

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 16:04:48 ; Search time 28.7 Seconds

(without alignments)
429.981 Million cell updates/sec

Title: US-08-699-716a-2

Sequence: 1 MGHNNHHHHHHSSGHIDDD.....RRIQKDYVWRLDTSCK 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2476.5	92.9	501	1 W01044	Y. pestis Fl/V ant
2	2382.5	89.4	480	1 W01045	Y. pestis Fl/V ant
3	1641.5	61.6	329	1 R79961	Partial lcrv (V an
4	1641.5	61.6	329	1 W01040	Y. pestis V antige
5	1641.5	61.6	329	1 W01041	Y. pestis V antige
6	1638	61.4	329	1 R79962	Partial lcrv (V an
7	850	31.9	170	1 R76528	Yersinia pestis ca
8	850	31.9	170	1 W01043	Y. pestis Fl antige
9	850	31.9	170	1 W59782	Amino acid sequenc
10	850	31.9	170	1 W59783	Amino acid sequenc
11	755	28.3	149	1 W59788	Nucleotide sequenc
12	755	28.3	150	1 W59787	Amino acid sequenc
13	755	28.3	151	1 R76526	Yersinia pestis ca
14	755	28.3	151	1 W01042	Y. pestis Fl antige
15	753	28.2	151	1 R76527	Yersinia pestis ca
16	753	25.2	192	1 W59785	Amino acid sequenc
17	578	21.7	171	1 W59786	Amino acid sequenc
18	179.5	6.7	451	1 W68398	Clostridium botuli
19	174.5	6.5	472	1 W68393	Clostridium botuli
20	170.5	6.4	472	1 W68394	Clostridium botuli
21	164.5	6.2	1254	1 W24575	Merozoite apical-e
22	164.5	6.2	1254	1 W19779	Polyhistidine-ente
23	159	6.0	473	1 W68400	Clostridium botuli
24	159	6.0	473	1 W68399	Clostridium botuli
25	153	5.7	448	1 W68397	Clostridium botuli
26	147	5.5	462	1 W68397	Clostridium botuli
27	145	5.4	966	1 R76065	Yeast MSH2 protein
28	144.5	5.4	672	1 R31216	Penicillin binding
29	144	5.4	462	1 R95009	Type A neurotoxin
30	144	5.4	462	1 W68390	Clostridium botuli
31	140.5	5.3	452	1 W68396	Clostridium botuli
32	139.5	5.2	423	1 W68098	Chlamydomonas rein
33	136.5	5.1	451	1 W68395	Clostridium botuli

34	132	5.0	53	1 W36986	HFRT peptide. Meth
35	132	5.0	54	1 W36990	HFRT peptide. Meth
36	130.5	4.9	147	1 R10534	Human 160kD mediat
37	128	4.8	687	1 W41586	Truncated restin p
38	125.5	4.7	799	1 W13126	Full length restin p
39	125.5	4.7	799	1 W25639	Full length restin p
40	124.5	4.7	455	1 Y00910	S. aureus ffn prot
41	124	4.7	573	1 W68202	M. catarrhalis str
42	123	4.6	1639	1 W54145	P. falciparum synt
43	123	4.6	1654	1 P50777	Sequence of the p1
44	123	4.6	2285	1 W98149	Bacillus subtilis
45	122	4.6	323	1 W71185	FP505 protein cont

ALIGNMENTS

RESULT	ID	Description	Score	Query Match	Length	ID	Description
1	W01044	W01044 standard; Protein; 501 AA.	92.98	Score 2476.5; DB 1; Length 501;			
AC	W01044	28-DEC-1996 (first entry)	98.68	Pred. NO. 2.1e-170;			
DE	Y. pestis Fl/V antigen fusion.			Matches 494; Conservative 1; Mismatches 3; Gaps 1;			
KW	Plaque; vaccine; genetic immunisation; V antigen; lcrv;						
OS	Chimeric Yersinia pestis strain GB;						
OS	Chimeric synthetic.						
FM	Key	location/Qualifiers					
FT	peptide	1..21					
FT		/label= Sig-peptide					
FT		/note= "Fl antigen signal peptide"					
FT	protein	22..170					
FT		/label= Fl antigen					
FT		/note= "mature Fl antigen"					
FT	peptide	171..176					
FT		/label= linker					
FT		/note= "6-amino acid peptide linker"					
FT	protein	176..501					
FT		/label= V antigen					
FT		/note= "mature V antigen"					
PD	W09628551-A1.						
PD	19-SEP-1996.						
PF	13-MAR-1996; G00571.						
PR	13-MAR-1995; GB-005059.						
PR	15-SEP-1995; GB-018946.						
PR	05-DEC-1995; GB-024825.						
PA	(MINA) UK SEC FOR DEFENCE.						
PI	Bennett AM, Leary SEC, Oyston PCF, Tiltall RW, Williamson ED;						
DR	WPI; 96-433824/43.						
DR	N-PSDB; T38249.						
PT	Yersinia pestis - useful in vaccine for protection against plague						
PS	epitopic parts - useful in vaccine for protection against plague						
CC	Example 3; Page 65-69; 98pp; English.						
CC	A fusion protein (W01044) comprises the Fl antigen (see also						
CC	W01043) and V antigen (see also W01041) of Yersinia pestis joined						
CC	by a linker that allows each protein to attain its conformational						
CC	state. It is the product of a gene fusion (T38249) obtd. by PCR						
CC	amplification of Y. pestis DNA. Fl/V fusion protein can be						
CC	expressed by gut-colonising organism transformants, to induce						
CC	an immune response against Y. pestis, the causative organism of						
CC	plague.						
SC	Sequence 501 AA;						
OY	24 MKKSSVIAALFETIPANADLTASTATATVEPARITLYKESAPITMDNGINDT 83						
DB	1 MKKSSVIAALFETIPANADLTASTATATVEPARITLYKESAPITMDNGINDT 60						
OY	84 ELAVGLTLLGGYKTKGTSTSVNFTDAAGDPMYLTFTSODGNHNOFTTKVIGKSRDPDIS 143						

Db 61 ELIVGLITLGGYTGTTSTSVNTDAGDPMTFTSODGNHQTFTKVLGKDSRDPDIS 120
 QY 144 PKVNGENLVGDVVLATGSDFFVRSIGSGKGLAAGKYDAVTYVSNQ---EFMIRAY 200
 Db 121 PKVNGENLVGDVVLATGSDFFVRSIGSGKGLAAGKYDAVTYVSNQSTIEGRIRAY 180
 QY 201 EONPOHFIEDLEKVRVQOLGHSVLEELVQVVKDKNDISIKYPRKDSVFANRVIT 260
 Db 181 EONPOHFIEDLEKVRVQOLGHSVLEELVQVVKDKNDISIKYPRKDSVFANRVIT 240
 QY 261 DDELLKTLIAYFLPEDTILKGGHYNOLONGIKRVEELESSPNTOMELRAFMAYMHS 320
 Db 241 DDELLKTLIAYFLPEDTILKGGHYNOLONGIKRVEELESSPNTOMELRAFMAYMHS 300
 QY 321 LTRADRIDDDILKIVDSMNHGDAKSLREELAEITAEIKIYSVIAEINKHLSSSGTIN 380
 Db 301 LTRADRIDDDILKIVDSMNHGDAKSLREELAEITAEIKIYSVIAEINKHLSSSGTIN 360
 QY 381 IHDKSINLMDKNLYGTDEEIFKASAEYKLEKMPOTTIOVDGSEKKIYSINDFLGSENK 440
 Db 361 IHDESNLMDKNLYGTDEEIFKASAEYKLEKMPOTTIOVDGSEKKIYSINDFLGSENK 420
 QY 441 RFGALGNLKNKSYNKNNDNELSHFATGSDKSRPLNDVSOQKTOLSDITSRNSAIEAL 500
 Db 421 RFGALGNLKNKSYNKNNDNELSHFATGSDKSRPLNDVSOQKTOLSDITSRNSAIEAL 480
 QY 501 NREIQRYDSVMORLDDTSGK 521
 Db 481 NREIQRYDSVMORLDDTSGK 501

RESULT 2

ID W01045 standard; Protein: 480 AA.
 AC W01045:
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1/V antigen fusion.
 KW Plaque; vaccine; genetic immunisation; V antigen; lcrV;
 KM F1 antigen; cat1.
 OS Chimeric Yersinia pestis strain GB.
 FH Chimeric synthetic.
 FT Key location/Qualifiers
 FT protein 1..149
 FT /label= F1 antigen
 FT /note= "mature F1 antigen"
 FT peptide 150..155
 FT /label= linker
 FT /note= "6-amino acid peptide linker"
 FT protein 156..480
 FT /label= V antigen
 FT /note= "mature V antigen"
 FT W09628551-A1.
 PD 19-SEP-1996.
 PF 13-MAR-1996: G00571.
 PR 13-MAR-1995: GB-005059.
 PR 15-SEP-1995: GB-018946.
 PR 05-DEC-1995: GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED.
 DR WPI: 96-433824/43.
 DR N-PSDB: T38256.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS disclosure: Page 51-55; 98pp; English.
 CC A fusion protein (W01045) comprises the F1 antigen (see also
 CC W01042) and V antigen (see also W01041) of Yersinia pestis joined
 CC by a linker that allows each protein to attain its conformational
 CC state. It is the product of a gene fusion (T38256) obt. by PCR
 CC amplification of Y. pestis DNA. F1/V fusion protein can be
 CC expressed by gut-colonising organism transformants, to induce
 CC an immune response against Y. pestis, the causative organism of
 CC plague.

SQ Sequence 480 AA:

Query Match 89.4%; Score 2382.5; DB 1; Length 480;
 Best local similarity 98.5%; Pred. No. 1,1e-163;
 Matches 473; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 45 ADITASTATATLVEPARITLTKKCAPITIMDNGIDELLVGLTLGGYTGTTSTSV 104
 Db 1 ADITASTATATLVEPARITLTKKCAPITIMDNGIDELLVGLTLGGYTGTTSTSV 60
 QY 105 NPTDAAGDPMYLTFTSODGNHQTFTKVLGKDSRDPDISPKVNGENLVGDVVLATGSD 164
 Db 61 NPTDAAGDPMYLTFTSODGNHQTFTKVLGKDSRDPDISPKVNGENLVGDVVLATGSD 120
 QY 165 FVFRSISGSGKGLAAGKYDAVTYVSNQ---EFMIRAYEONPOHFIEDLEKVRVQOLTG 221
 Db 121 FVFRSISGSGKGLAAGKYDAVTYVSNQSTIEGRIRAYEONPOHFIEDLEKVRVQOLTG 180
 QY 222 HGSSVLEELVQVVKDKNDISIKYPRKDSVFANRVITDIELLKLILAYFLPEDTILK 281
 Db 181 HGSSVLEELVQVVKDKNDISIKYPRKDSVFANRVITDIELLKLILAYFLPEDTILK 240
 QY 282 GGHYNQLONGIKRVEELESSPNTOMELRAFMAYMHSLTRADRIDDDILKIVDSMNH 341
 Db 241 GGHYNQLONGIKRVEELESSPNTOMELRAFMAYMHSLTRADRIDDDILKIVDSMNH 300
 QY 342 GDARSKLREELAEITAEIKIYSVIAEINKHLSSSGTINIHDKSINLMDKNLYGTDEI 401
 Db 301 GDARSKLREELAEITAEIKIYSVIAEINKHLSSSGTINIHDKSINLMDKNLYGTDEI 360
 QY 402 FRASAEYKLEKMPOTTIOVDGSEKKIYSINDFLGSENKRTGALNKNKSYNKNNDNEL 461
 Db 361 FRASAEYKLEKMPOTTIOVDGSEKKIYSINDFLGSENKRTGALNKNKSYNKNNDNEL 420
 QY 462 SHFATGSDKSRPLNDVSOQKTOLSDITSRNSAIEALNREIQRYDSVMORLDDTSGK 521
 Db 421 SHFATGSDKSRPLNDVSOQKTOLSDITSRNSAIEALNREIQRYDSVMORLDDTSGK 480

RESULT 3

ID R79961 standard; Protein: 329 AA.
 AC R79961:
 DT 18-APR-1996 (first entry)
 DE Partial lcrV (V antigen) of Y. pestis.
 KW lcrV; V antigen; virulence; plaque; vaccine; epitope.
 OS Yersinia pestis.
 FT W095244/5-A1.
 PD 14-SEP-1995.
 PF 06-MAR-1995: G00481.
 PR 08-MAR-1994: GB-004577.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Leary SEC, Titball RW, Williamson ED, Leary SE.
 DR WPI: 95-328268/42.
 DR N-PSDB: T04222.
 PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
 PT oral or parenteral vaccines for protection against plague
 PS Claim 6: Page 11-13; 25pp; English.
 CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrV) encoding
 CC all or a protective epitopic part of the mature V protein of Yersinia
 CC pestis. The protein was expressed as a fusion protein with maltose
 CC binding protein or glutathione-S-transferase in 3 different plasmid
 CC vectors. Y. pestis is the highly virulent causative organism of plague
 CC in a wide range of animals, including man. The V antigen (lcrV) is an
 CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
 CC The V antigen is postulated to act as a virulence antigen, and
 CC transformed microorganisms contg. recombinant DNA encoding a V antigen
 CC protein/peptide are useful in vaccines to protect against plague.
 SO Sequence 329 AA;

Query Match

61.6%; Score 1641.5; DB 1; Length 329;

Best Local Similarity 99.4%; Pred. No. 1.2e-110;
Matches 326; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 194 EFMRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 253
| | | | |
DB 3 EF-IRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 61
| | | | |
QY 254 FANRVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 313
| | | | |
DB 62 FANRVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 121
| | | | |
QY 314 MAVHFSFLTADRIDDDILKIVYDSNMHGDARSKLREELAEITAEIKIYSVIOAEINKHL 373
| | | | |
DB 122 MAVHFSFLTADRIDDDILKIVYDSNMHGDARSKLREELAEITAEIKIYSVIOAEINKHL 181
| | | | |
QY 374 SSGGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD 433
| | | | |
DB 182 SSGGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD 241
| | | | |
QY 434 FLGSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVSOQTQOLSDITSRF 493
| | | | |
DB 242 FLGSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVSOQTQOLSDITSRF 301
| | | | |
QY 494 NSAIETALNRFIOKYDSVMQRLDDTSCK 521
| | | | |
DB 302 NSAIETALNRFIOKYDSVMQRLDDTSCK 329
| | | | |

RESULT 4
ID W01040 standard; Protein; 329 AA.
AC W01040;
DT 28-DEC-1996 (first entry)
DE Y. pestis V antigen.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KM Yersinia pestis strain GB.
OS Yersinia pestis strain GB.
FH Key location/Qualifiers
FT peptide 1..4 /note- "vector-encoded peptide"
PN W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tlball RW, Williamson ED;
DR WPI; 96-433824/43.
DR N-PSDB; T38242.
PT Yersinia pestis V antigen and F1 antigen or their protective
PS epitopic parts - useful in vaccine for protection against plague
PS Disclosure; Page 25-28; 98pp; English.
CC Yersinia pestis V antigen (W01040) is capable of evoking protective
CC immune responses in animals. A combined vaccine of V and F1
CC antigens (see also W01042) can at least match the protection
CC afforded by live attenuated EV76 vaccine without any of the hazards
CC that have kept the EV vaccine from general use. The V antigen is
CC produced using the lcrV gene (see also T38242) obtd. from Y.
CC pests by PCR amplification. It can also be prepd. as a fusion with
CC F1 antigen (see also W01044-45) and expressed by attenuated Arca or
CC C Salmonella typhi as a live vaccine for long-term protection
CC against plague. Expression by gut-colonising bacterial transformants
CC produces a protective response against Y. pestis.
SQ Sequence 329 AA;

Query Match 61.6%; Score 1641.5; DB 1; Length 329;
Best Local Similarity 99.4%; Pred. No. 1.2e-110;
Matches 326; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 194 EFMRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 253
| | | | |
DB 3 EF-IRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 61
| | | | |
QY 254 FANRVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 313
| | | | |
DB 62 FANRVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 121
| | | | |
QY 314 MAVHFSFLTADRIDDDILKIVYDSNMHGDARSKLREELAEITAEIKIYSVIOAEINKHL 373
| | | | |
DB 122 MAVHFSFLTADRIDDDILKIVYDSNMHGDARSKLREELAEITAEIKIYSVIOAEINKHL 181
| | | | |
QY 374 SSGGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD 433
| | | | |
DB 182 SSGGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD 241
| | | | |
QY 434 FLGSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVSOQTQOLSDITSRF 493
| | | | |
DB 242 FLGSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVSOQTQOLSDITSRF 301
| | | | |
QY 494 NSAIETALNRFIOKYDSVMQRLDDTSCK 521
| | | | |
DB 302 NSAIETALNRFIOKYDSVMQRLDDTSCK 329
| | | | |

Query Match 61.6%; Score 1641.5; DB 1; Length 329;
Best Local Similarity 99.4%; Pred. No. 1.2e-110;
Matches 326; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 194 EFMRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 253
| | | | |
DB 3 EF-IRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 61
| | | | |
QY 254 FANRVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 313
| | | | |
DB 62 FANRVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 121
| | | | |
QY 314 MAVHFSFLTADRIDDDILKIVYDSNMHGDARSKLREELAEITAEIKIYSVIOAEINKHL 373
| | | | |
DB 122 MAVHFSFLTADRIDDDILKIVYDSNMHGDARSKLREELAEITAEIKIYSVIOAEINKHL 181
| | | | |
QY 374 SSGGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD 433
| | | | |
DB 182 SSGGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD 241
| | | | |
QY 434 FLGSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVSOQTQOLSDITSRF 493
| | | | |
DB 242 FLGSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVSOQTQOLSDITSRF 301
| | | | |
QY 494 NSAIETALNRFIOKYDSVMQRLDDTSCK 521
| | | | |
DB 302 NSAIETALNRFIOKYDSVMQRLDDTSCK 329
| | | | |

DB 3 EF-IRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 61
| | | | |
QY 254 FANRVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 313
| | | | |
DB 62 FANRVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 121
| | | | |
QY 314 MAVHFSFLTADRIDDDILKIVYDSNMHGDARSKLREELAEITAEIKIYSVIOAEINKHL 373
| | | | |
DB 122 MAVHFSFLTADRIDDDILKIVYDSNMHGDARSKLREELAEITAEIKIYSVIOAEINKHL 181
| | | | |
QY 374 SSGGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD 433
| | | | |
DB 182 SSGGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD 241
| | | | |
QY 434 FLGSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVSOQTQOLSDITSRF 493
| | | | |
DB 242 FLGSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVSOQTQOLSDITSRF 301
| | | | |
QY 494 NSAIETALNRFIOKYDSVMQRLDDTSCK 521
| | | | |
DB 302 NSAIETALNRFIOKYDSVMQRLDDTSCK 329
| | | | |

RESULT 5
ID W01041 standard; Protein; 329 AA.
AC W01041;
DT 28-DEC-1996 (first entry)
DE Y. pestis V antigen.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KM Yersinia pestis strain GB.
OS Yersinia pestis strain GB.
FH Key location/Qualifiers
FT peptide 1..4 /note- "vector-encoded peptide"
PN W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tlball RW, Williamson ED;
DR WPI; 96-433824/43.
DR N-PSDB; T38243.
PT Yersinia pestis V antigen and F1 antigen or their protective
PS epitopic parts - useful in vaccine for protection against plague
PS Disclosure; Page 32-35; 98pp; English.
CC Yersinia pestis V antigen (W01041) is capable of evoking protective
CC immune responses in animals. A combined vaccine of V and F1
CC antigens (see also W01042) can at least match the protection
CC afforded by live attenuated EV76 vaccine without any of the hazards
CC that have kept the EV vaccine from general use. The V antigen is
CC produced using the lcrV gene (see also T38243) obtd. from Y. pestis
CC by PCR amplification. It can also be prepd. as a fusion with F1
CC antigen (see also W01044-45) and expressed by attenuated Arca or C
CC Salmonella typhi and gut-colonising bacteria for protection against
CC plague.
SQ Sequence 329 AA;

Query Match 61.6%; Score 1641; DB 1; Length 329;
Best Local Similarity 99.7%; Pred. No. 1.3e-110;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 197 IRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 256
| | | | |
DB 5 IRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 64
| | | | |
QY 257 RVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 316
| | | | |
DB 65 RVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 124
| | | | |

Query Match 61.6%; Score 1641; DB 1; Length 329;
Best Local Similarity 99.7%; Pred. No. 1.3e-110;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 197 IRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 256
| | | | |
DB 5 IRAYEQNPQHFIEDLEKRVYEQLTGHSVLEELVQVKNDNISIKYDPRKSEV 64
| | | | |
QY 257 RVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 316
| | | | |
DB 65 RVITDDIELLKKILAYFLPEDAILKGHYDNOQNGIKRVEFLESSPNTQWELRAF 124
| | | | |

QY	317	MHFSLTARIDDDILKVIYDSMMHHHDASKLREELAEELTAELKITSYVQOAEINKLHSSS	376
Db	125	MHFSLTARIDDDILKVIYDSMMHHHDASKLREELAEELTAELKITSYVQOAEINKLHSSS	184
QY	377	GTINIHDSINIMLMDKNLYGTDEEIFKSAEKKILEKMPOTTIOVGSEKKIYISIDFLG	436
Db	185	GTINIHDSINIMLMDKNLYGTDEEIFKSAEKKILEKMPOTTIOVGSEKKIYISIDFLG	244
QY	437	SENKRTGALGNLKNKSNYKNNNELSHFATGSDKSRPLNDLYSQKTTQSLDITSRPNNA	496
Db	245	SENKRTGALGNLKNKSNYKNNNELSHFATGSDKSRPLNDLYSQKTTQSLDITSRPNNA	304
QY	497	IEALNRFIOKDYDVAQRLDDTSGK 521	
Db	305	IEALNRFIOKDYDVAQRLDDTSGK 329	
RESULT	6		
ID	R79962		
AC	R79962 standard; Protein: 329 AA.		
DT	18-APR-1996 (first entry)		
DE	Partial LcrV (V antigen) of Y. pestis.		
KW	LcrV; V antigen; virulence; plague; vaccine; epitope.		
FN	Yersinia pestis.		
PN	W09524475-A1.		
PD	14-SEP-1995		
PF	06-MAR-1995; G00481.		
PR	08-MAR-1994; GB-004577.		
PA	(MUNA) UK SEC FOR DEFENCE.		
PI	Leary SEC, Tibball RW, Williamson ED, Leary SE;		
DR	WPI: 95-338268/42.		
DR	N-PSDB: T04223.		
PT	Recombinant DNA expressing Yersinia pestis V antigen - useful in		
PT	oral or parenteral vaccines for protection against plague		
PS	Claim 6; Page 15-16; 25pp; English.		
CC	R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrV) encoding		
CC	all or a protective epitopic part of the mature V protein of Yersinia		
CC	pestis. The protein was expressed as a fusion protein with maltose		
CC	binding protein or glutathione-S-transferase in 3 different plasmid		
CC	vectors. Y. pestis is the highly virulent causative organism of plague		
CC	in a wide range of animals, including man. The V antigen (LcrV) is an		
CC	unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb Lcr plasmid.		
CC	The V antigen is postulated to act as a virulence antigen, and		
CC	transformed microorganisms conty. recombinant DNA encoding a V antigen		
CC	protein/peptide are useful in vaccines to protect against plague.		
SO	Sequence 329 AA.		
Query Match	61.4%; Score 1638; DB 1; Length 329;		
Best Local Similarity	99.4%; Pred. No. 2, 1e-110;		
Matches	323; Conservative 1; Mismatches 1; Indels 0; Gaps 0		
QY	197	IRAYQONPOHFLIEDLEKVAVEQDTGHSVLEELVOLVAKDNIDISIKYDPKROSEVFAN	256
Db	5	IRAYQONPOHFLIEDLEKVAVEQDTGHSVLEELVOLVAKDNIDISIKYDPKROSEVFAN	64
QY	257	RYITDIDELKKILAVFLPEDIITLKGHDNOLONGIKRVEKLESSPTMQWELRAFMAY	316
Db	65	RYITDIDELKKILAVFLPEDIITLKGHDNOLONGIKRVEKLESSPTMQWELRAFMAY	124
QY	317	MHFSLTARIDDDILKVIYDSMMHHHDASKLREELAEELTAELKITSYVQOAEINKLHSSS	376
Db	125	MHFSLTARIDDDILKVIYDSMMHHHDASKLREELAEELTAELKITSYVQOAEINKLHSSS	184
QY	377	GTINIHDSINIMLMDKNLYGTDEEIFKSAEKKILEKMPOTTIOVGSEKKIYISIDFLG	436
Db	185	GTINIHDSINIMLMDKNLYGTDEEIFKSAEKKILEKMPOTTIOVGSEKKIYISIDFLG	244
QY	437	SENKRTGALGNLKNKSNYKNNNELSHFATGSDKSRPLNDLYSQKTTQSLDITSRPNNA	496
Db	245	SENKRTGALGNLKNKSNYKNNNELSHFATGSDKSRPLNDLYSQKTTQSLDITSRPNNA	304

```

Yy      497    IEALNRFIOKDYVNMQRLLDDTSCK   521
          |||
Db      305    IEALNRFIOKIDYSVMQRLDDTSGK   329

RESULT      7
R76528
ID      R76528 standard; Protein; 170 AA.
AC      R76528;
DT      17-DEC-1995 (first entry)
De      Versinia pestis cafi (F1) antigen.
KW      Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
        bubonic plague; pneumonic plague.
OS      Versinia pestis.
PN      W09518231-A1.
PD      06-JUL-1995.
PE      23-DEC-1994; G02818.
PR      24-DEC-1993; GB-026425.
PA      (MINA.) UK SEC FOR DEFENCE.
PI      Howells A., Deary SEC, Oyston PCF, Titchall RW, Williamson ED;
DR      WPJ.; 95-246396/32.
DR      N-PDSB; Q92819.
PT      DNA constructs capable of transforming microorganisms - which can be
PT      used as live or attenuated vaccines which induce an immune response,
PT      against Versinia pestis, at mucosal surfaces.
PS      Disclosure: Page 20; 2/PP; English.
CC      The sequence represents the Y. pestis cafi (F1) antigen expressed
CC      from plasmid pFORFLD. The DNA construct can be used to transform
CC      human or animal gut colonizing microorganisms, specifically
CC      attenuated Salmonella typhimurium or Salmonella typhi. The
CC      transformed microorganisms can be used as live/attenuated vaccines
CC      which induce immune responses at mucosal surfaces. The vaccines
CC      provide protection against infection with Y. pestis, and are
CC      parenterally and orally active vaccines offering protection
CC      against bubonic and pneumonic plague.
SO      Sequence 170 AA;

Query Match      31.9%; Score 850; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 3.6e-54;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy      24      MKKISSVAIALFGITATANADLTAATAATLVLEPARITYTYEGAPITIMDNINDT   83
          |||
Db      1      MKKISSVAIALFGITATANADLTAATAATLVLEPARITYTYEGAPITIMDNINDT   60

Yy      84      ELVTTLTGKGKGTSTSVNFETGAAGPWLFEETSDGNNHQTIVKIGDSRFDIS   143
          |||
Db      61      ELVTTLTGKGKGTSTSVNFETGAAGPWLFTTSQGNHQTIVKIGDSRFDIS   120

Yy      144     PRVNGENLVGDVDVLATGSODEFFVRSIGSKGSKIAGKYTDVAVTVSNQ   193
          |||
Db      121     PRVNGENLVGDVDVLATGSODEFFVRSIGSKGSKLAGKRTDAVTVTVSNQ   170

RESULT      8
W01043
ID      W01043 standard; Protein; 170 AA.
AC      W01043;
DT      28-DEC-1996 (first entry)
De      Y. pestis F1 antigen (including signal peptide).
KW      Plague; vaccine; genetic immunisation; F1 antigen; cafi;
        V antigen.
OS      Versinia pestis strain GB.
FH      Key location/Qualifiers
FT      peptide 1..21
FN      W09628851-A1.
PN      13-SEP-1996.
PD      13-MAR-1996; G00571.
PE      13-MAR-1995; GB-005059.
PR      13-SEP-1995; GB-018946.
PA      05-DEC-1995; GB-024825.

```

PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tlballi RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR N-PSDB: T38248.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Example 2; Page 61-62; 98pp; English.
 CC The F1 antigen (W01043), including the signal peptide, of Yersinia
 CC pestis was produced from a DNA sequence (T38248) obtd. by PCR
 CC amplification (see also T38257-58) of Y. pestis DNA. Expression
 CC of the F1 antigen (see also W01042) by gut-colonising organisms in
 CC the form of live vaccines can be used to protect an animal,
 CC including humans, against plague.
 SQ Sequence 170 AA;

Query Match 31.9%; Score 850; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 3.6e-54;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAIALFGTITANADLASTATATLVEPARITLYTEGAPITIMNGNIDT 83
 DB 1 MKKISSVIAIALFGTITANADLASTATATLVEPARITLYTEGAPITIMNGNIDT 60
 QY 84 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 143
 DB 61 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 120
 QY 144 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 193
 DB 121 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 170

RESULT 9
 W59782
 ID W59782 standard; Protein; 170 AA.
 AC W59782;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pyF1sec170.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 DR N-PSDB: V41594.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 52; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 170 AA;

Query Match 31.9%; Score 850; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 3.6e-54;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAIALFGTITANADLASTATATLVEPARITLYTEGAPITIMNGNIDT 83
 DB 1 MKKISSVIAIALFGTITANADLASTATATLVEPARITLYTEGAPITIMNGNIDT 60
 QY 84 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 143
 DB 61 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 120
 QY 144 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 193

DB 121 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 170

RESULT 10
 W59783
 ID W59783 standard; Protein; 170 AA.
 AC W59783;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen nyF1sec510.
 KM F1 antigen; plasmid; vaccine; plague; ds.
 OS Yersinia pestis.
 FH Key Location/Qualifiers
 FT CDS 17..532
 FT /tag- a
 FT /product- "F1 antigen"
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 DR N-PSDB: V41596.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Disclosure; Pages 53-54; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 170 AA;

Query Match 31.9%; Score 850; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 3.6e-54;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAIALFGTITANADLASTATATLVEPARITLYTEGAPITIMNGNIDT 83
 DB 1 MKKISSVIAIALFGTITANADLASTATATLVEPARITLYTEGAPITIMNGNIDT 60
 QY 84 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 143
 DB 61 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 120
 QY 144 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 193
 DB 121 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 170

RESULT 11
 W59788
 ID W59788 standard; Protein; 149 AA.
 AC W59788;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen pyF1mat149.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 DR N-PSDB: V41609.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 63; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to

CC Produce recombinant antigens, especially Yersinia pestis antigens.
 CC Protecting an animal from contracting plague.
 CC Sequence 149 AA;

Query Match 28.3%; Score 755; DB 1; Length 149;
 Best Local Similarity 100.0%; Pred. No. 2e-47;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 ADLTASTATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSV 104
 DB 1 ADLTASTATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSV 60
 OY 105 NETDAAGDPMYLTFTSODGNHOFITKVGKSRDPDISPKVNGENLVGDDVVLATGSOD 164
 DB 61 NETDAAGDPMYLTFTSODGNHOFITKVGKSRDPDISPKVNGENLVGDDVVLATGSOD 120
 OY 165 FFVRSIGSGKGLAAGKYTDAVTVYSNQ 193
 DB 121 FFVRSIGSGKGLAAGKYTDAVTVYSNQ 149

RESULT 12

W59787
 ID W59787 standard; Protein: 150 AA.

AC W59787.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPFmat150.
 KW F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HEK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 DR N-PSDB: V41600.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 60; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 150 AA;

Query Match 28.3%; Score 755; DB 1; Length 150;
 Best Local Similarity 100.0%; Pred. No. 2e-47;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 ADLTASTATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSV 104
 DB 2 ADLTASTATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSV 61
 OY 105 NETDAAGDPMYLTFTSODGNHOFITKVGKSRDPDISPKVNGENLVGDDVVLATGSOD 164
 DB 62 NETDAAGDPMYLTFTSODGNHOFITKVGKSRDPDISPKVNGENLVGDDVVLATGSOD 121
 OY 165 FFVRSIGSGKGLAAGKYTDAVTVYSNQ 193
 DB 122 FFVRSIGSGKGLAAGKYTDAVTVYSNQ 150

RESULT 13

R76526
 ID R76526 standard; Protein: 151 AA.
 AC R76526;
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis cafi (F1) antigen.

KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 KW bubonic plague; pneumonic plague.
 OS Yersinia pestis.
 PN W09518231-A1.
 PD 06-JUL-1995.
 PF 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
 DR WPI: 95-246396/32.
 DR N-PSDB: G92817.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure; Page 16; 27pp; English.
 CC The sequence represents the Y. pestis cafi (F1) antigen expressed
 CC from plasmid pFGAL2s. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated Salmonella typhimurium or Salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 151 AA;

Query Match 28.3%; Score 755; DB 1; Length 151;
 Best Local Similarity 98.0%; Pred. No. 2e-47;
 Matches 148; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 43 NADLTASTATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSV 102
 DB 1 SSADLTASTATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSV 60
 OY 103 SYNFTDAAGDPMYLTFTSODGNHOFITKVGKSRDPDISPKVNGENLVGDDVVLATGS 162
 DB 61 SYNFTDAAGDPMYLTFTSODGNHOFITKVGKSRDPDISPKVNGENLVGDDVVLATGS 120
 OY 163 ODFFVRSIGSGKGLAAGKYTDAVTVYSNQ 193
 DB 121 ODFFVRSIGSGKGLAAGKYTDAVTVYSNQ 151

RESULT 14

W01042
 ID W01042 standard; Protein: 151 AA.

AC W01042.
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1 antigen.
 KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;
 KW V antigen.
 OS Yersinia pestis strain GB.
 PN W09628551-A1.
 PD 19-SEP-1996.
 PF 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MTNA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR N-PSDB: T38244.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Disclosure; Page 43-45; 98pp; English.
 CC Yersinia pestis F1 antigen (W01042) is capable of evoking protective
 CC immune responses in animals. A combined vaccine of F1 and V
 CC antigens (see also W01040-41) can at least match the protection
 CC afforded by live attenuated EV76 vaccine without any of the hazards
 CC that have kept the EV vaccine from general use. The F1 antigen is
 CC produced using the cafi gene (see also T38244) obtd. from Y.
 CC pestis by PCR amplification. It can also be prepd. as a fusion with

CC V antigen (see also W01044-45) and expressed by attenuated AroA or
 CC C Salmonella typhi as a live vaccine for long-term protection
 CC against plague. Expression by gut-colonising bacterial transformants
 CC produces a protective response against Y. pestis.
 SO Sequence 151 AA;

Search completed: August 22, 2000, 16:46:38
 Job time: 2510 sec

Query Match 28.3%; Score 755; DB 1; Length 151;
 Best Local Similarity 98.0%; Pred. No. 2e-47;
 Matches 148; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 43 NAADLTATATATATVEPARITLYKEGAPITIMDNGNIDTELVGTLTGKYKTGTST 102
 DB 1 SSADLTATATATATVEPARITLYKEGAPITIMDNGNIDTELVGTLTGKYKTGTST 60
 QY 103 SVNFTDAGDPMYLTFTSODGNHOFITKVGKDSRDPDISPKVNGENLVGDVYLATGS 162
 DB 61 SVNFTDAGDPMYLTFTSODGNHOFITKVGKDSRDPDISPKVNGENLVGDVYLATGS 120
 QY 163 ODFEVRISIGSKGKLAGKYTDATVTVSNQ 193
 DB 121 ODFEVRISIGSKGKLAGKYTDATVTVSNQ 151

RESULT 15

R76527
 ID R76527 standard; Protein; 151 AA.
 AC R76527;
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis ca1 (F1) antigen.
 KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 KW bubonic plague; pneumonic plague.
 OS Yersinia pestis.
 PN W09518231-A1.
 PD 06-JUL-1995.
 PF 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
 DR WPI: 95-246396/32.
 DR N-PSDB; 092818.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure: Page 18; 27pp; English.
 CC The sequence represents the Y. pestis ca1 (F1) antigen expressed
 CC from plasmid pF51G3a. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated Salmonella typhimurium or Salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SO Sequence 151 AA;

Query Match 28.2%; Score 753; DB 1; Length 151;
 Best Local Similarity 99.3%; Pred. No. 2.8e-47;
 Matches 148; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 45 ADLTATATATATVEPARITLYKEGAPITIMDNGNIDTELVGTLTGKYKTGTSTSV 104
 DB 3 ADLTATATATATVEPARITLYKEGAPITIMDNGNIDTELVGTLTGKYKTGTSTSV 62
 QY 105 NETDAAGDPMYLTFTSODGNHOFITKVGKDSRDPDISPKVNGENLVGDVYLATGSOD 164
 DB 63 NETDAAGDPMYLTFTSODGNHOFITKVGKDSRDPDISPKVNGENLVGDVYLATGSOD 122
 QY 165 FFEVRISIGSKGKLAGKYTDATVTVSNQ 193
 DB 123 FFEVRISIGSKGKLAGKYTDATVTVSNQ 151

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 16:29:03 ; Search time 22.92 seconds
(without alignments)

348,451 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 2666
Sequence: 1 MGHNNHHHHHHSSGHIDD.....RFIOKYDSVWRLDDTSGR 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 152396 seqs, 15329161 residues

Total number of hits satisfying chosen parameters: 152396

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2480.5	93.0	501	2	US-08-913-477-23
2	2378.5	89.2	480	2	US-08-913-477-17
3	1641.5	61.6	329	2	US-08-913-477-2
4	1641	61.6	329	2	US-08-913-477-4
5	844	31.7	170	2	US-08-913-477-21
6	755	28.3	151	2	US-08-913-477-11
7	159	6.0	24	3	US-08-780-496-8
8	144	5.4	462	1	US-08-480-604A-26
9	144	5.4	462	2	US-08-405-496A-26
10	132	5.0	53	2	US-08-651-818A-19
11	132	5.0	54	2	US-08-651-818A-23
12	125.5	4.7	799	1	US-08-188-228-42
13	125.5	4.7	799	1	US-08-332-638-42
14	124	4.6	1683	3	US-08-735-587-13
15	122	4.6	23	1	US-08-480-604A-24
16	122	4.6	23	1	US-08-405-496A-24
17	122	4.6	323	1	US-08-591-989-5
18	119.5	4.5	793	1	US-08-188-228-54
19	119.5	4.5	793	1	US-08-332-643-48
20	119.5	4.5	793	1	US-08-332-638-54
21	119	4.5	21	2	US-08-651-818A-21
22	118.5	4.4	532	1	US-08-188-228-44
23	118.5	4.4	532	1	US-08-332-638-44
24	118	4.4	800	3	US-08-776-265-3
25	115	4.3	859	1	US-08-053-614-2
26	115	4.3	859	1	US-08-316-397B-2
27	115	4.3	859	2	US-09-034-306-2
28	115	4.3	859	4	PCT-US93-09782-2

29	115	4.3	1181	1	US-08-053-614-4	Sequence 4, Appl1
30	115	4.3	1181	1	US-08-316-397B-4	Sequence 4, Appl1
31	115	4.3	1181	2	US-09-034-306-4	Sequence 4, Appl1
32	115	4.3	1181	4	PCT-US93-09782-4	Sequence 4, Appl1
33	113.5	4.3	792	3	US-08-433-522A-8	Sequence 8, Appl1
34	113	4.2	1354	3	US-08-685-871-2	Sequence 2, Appl1
35	112	4.2	711	3	US-08-949-588-2	Sequence 2, Appl1
36	112	4.2	2329	4	US-08-755-587-16	Sequence 16, Appl1
37	112	4.2	2938	4	PCT-US94-00198-3	Sequence 3, Appl1
38	112	4.2	3111	2	US-08-460-309-4	Sequence 4, Appl1
39	112	4.2	3111	2	US-08-125-077-4	Sequence 4, Appl1
40	112	4.2	3418	2	US-08-639-501-2	Sequence 2, Appl1
41	112	4.2	3418	3	US-09-044-946-2	Sequence 2, Appl1
42	111.5	4.2	736	1	US-07-688-352C-24	Sequence 24, Appl1
43	111.5	4.2	736	2	US-08-474-379C-24	Sequence 24, Appl1
44	111.5	4.2	736	3	US-09-146-249A-24	Sequence 24, Appl1
45	111.5	4.2	736	4	PCT-US91-02714-23	Sequence 23, Appl1

ALIGNMENTS

RESULT 1
US-08-913-477-23
Sequence 23, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Gable Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid

TOPOLOGY: 1linear
MOLECULE TYPE: protein
US-08-913-477-23

Query Match 93.0%; Score 2480.5; DB 2; Length 501;
Best Local Similarity 98.8%; Pred. No. 3,7e-187;
Matches 495; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 24 MKRISSVIAIALEFGTATANAADLSTATAFLVEPARITLTKKGAPITINDNGNIDT 83
DB 1 MKRISSVIAIALEFGTATANAADLSTATAFLVEPARITLTKKGAPITINDNGNIDT 60
QY 84 ELIVGLTLTGKGTGTSTSVNFTDAAGDPMTLFTSQDGNHGFTHYVGNKSRPDIS 143
DB 61 ELIVGLTLTGKGTGTSTSVNFTDAAGDPMTLFTSQDGNHGFTHYVGNKSRPDIS 120
QY 144 PKYNGENLVGDDVYLATGSDPFVRSIGSKGKLAGKYDAVTWVSNO--EFMIRAY 200
DB 121 PKYNGENLVGDDVYLATGSDPFVRSIGSKGKLAGKYDAVTWVSNOGSIKGRIRAY 180
QY 201 EONPQHIEDLEKRVYEQTLGHSSVLEELVQVYKDKNIDISIKYPRKDSVFANRVIT 260
DB 181 EONPQHIEDLEKRVYEQTLGHSSVLEELVQVYKDKNIDISIKYPRKDSVFANRVIT 240
QY 261 DDEILKTLIAVFLPBDTILKGGHYDNOQNGIKRKEFELESSPNTOMELRAVMAVHFS 320
DB 241 DDEILKTLIAVFLPBDTILKGGHYDNOQNGIKRKEFELESSPNTOMELRAVMAVHFS 300
QY 321 LTPADRIDDLILKIVDSMNHGDAKSLRELAELTAELKIVSVIOAEIKRHLSSGGIN 380
DB 301 LTPADRIDDLILKIVDSMNHGDAKSLRELAELTAELKIVSVIOAEIKRHLSSGGIN 360
QY 381 IHDKSNLMDKNLYGTDEEIFKASAEYKILEKMPOTTIQVDSEKKIYSIKDFLSENK 440
DB 361 IHDKSNLMDKNLYGTDEEIFKASAEYKILEKMPOTTIQVDSEKKIYSIKDFLSENK 420
QY 441 RFGALNLNKSYSYNKDNNEISFATTCSDKSRPLNDVYSOKTQSLDITSRNSAIEAL 500
DB 421 RFGALNLNKSYSYNKDNNEISFATTCSDKSRPLNDVYSOKTQSLDITSRNSAIEAL 480
QY 501 NREIQKYSVMORLDDTSGK 521
DB 481 NREIQKYSVMORLDDTSGK 501

RESULT 2
US-08-913-477-17
Sequence 17, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: 1linear
MOLECULE TYPE: protein
US-08-913-477-17

Query Match 89.2%; Score 2378.5; DB 2; Length 480;
Best Local Similarity 98.5%; Pred. No. 3.6e-179;
Matches 473; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 45 ADLSTATAFLVEPARITLTKKGAPITINDNGNIDTELLVGLTLTGKGTGTSTSV 104
DB 1 ADLSTATAFLVEPARITLTKKGAPITINDNGNIDTELLVGLTLTGKGTGTSTSV 60
QY 105 NPTDAAGDPMTLFTSQDGNHGFTHYVGNKSRPDISPKYNGENLVGDDVYLATGSD 164
DB 61 NPTDAAGDPMTLFTSQDGNHGFTHYVGNKSRPDISPKYNGENLVGDDVYLATGSD 120
QY 165 FPFVRSIGSKGKLAGKYDAVTWVSNO--EFMIRAYEONPQHIEDLEKRVYEQTLG 221
DB 121 FPFVRSIGSKGKLAGKYDAVTWVSNOGSIKGRIRAYEONPQHIEDLEKRVYEQTLG 180
QY 222 HGSVLEELVQVYKDKNIDISIKYPRKDSVFANRVITDDEILKTLIAVFLPBDTILK 281
DB 181 HGSVLEELVQVYKDKNIDISIKYPRKDSVFANRVITDDEILKTLIAVFLPBDTILK 240
QY 282 GGHYDNOQNGIKRKEFELESSPNTOMELRAVMAVHFSLTPADRIDDLILKIVDSMNH 341
DB 241 GGHYDNOQNGIKRKEFELESSPNTOMELRAVMAVHFSLTPADRIDDLILKIVDSMNH 300
QY 342 GDAKSLRELAELTAELKIVSVIOAEIKRHLSSGGINHDKSNLMDKNLYGTDEEI 401
DB 301 GDAKSLRELAELTAELKIVSVIOAEIKRHLSSGGINHDKSNLMDKNLYGTDEEI 360
QY 402 FRASAEYKILEKMPOTTIQVDSEKKIYSIKDFLSENKRTGALNLNKSYSYNKDNNE 461
DB 361 FRASAEYKILEKMPOTTIQVDSEKKIYSIKDFLSENKRTGALNLNKSYSYNKDNNE 420
QY 462 SHFATTCSDKSRPLNDVYSOKTQSLDITSRNSAIEALNREIQKYSVMORLDDTSGK 521
DB 421 SHFATTCSDKSRPLNDVYSOKTQSLDITSRNSAIEALNREIQKYSVMORLDDTSGK 480

RESULT 3
US-08-913-477-2
Sequence 2, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.

APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-2

Query Match 61.6%; Score 1641.5; DB 2; Length 329;
Best Local Similarity 99.4%; Pred. No. 1.6e-121;
Matches 326; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 194 EFMRYAQNPOHFEDEKRVVEQLTGHSSVLEELVQVYKDKNDISIKYDPRKDESEV 253
DB 3 EF-IRAYEQNPOHFEDEKRVVEQLTGHSSVLEELVQVYKDKNDISIKYDPRKDESEV 61
OY 254 FANRYITDDIELKKILAYFLPEDTITLKGGHYDNOLONGIKRVKFELESSPNTOWELRAF 313
DB 62 FANRYITDDIELKKILAYFLPEDTITLKGGHYDNOLONGIKRVKFELESSPNTOWELRAF 121
OY 314 MAVMFSLTADRIDDDILKVIYDSNNHHGDARSKLREELAEITAEIKIYSVQAEINKHL 373
DB 122 MAVMFSLTADRIDDDILKVIYDSNNHHGDARSKLREELAEITAEIKIYSVQAEINKHL 181
OY 374 SSSGTTINIHDSINIMDKNLYGTDEEITFKASAEYKIIKEMPOTTIOVDGSEKKIYSIKD 433
DB 182 SSSGTTINIHDSINIMDKNLYGTDEEITFKASAEYKIIKEMPOTTIOVDGSEKKIYSIKD 241
OY 434 FLGSNKRKRGALGNLKNSTYKKNNDNLSHFATITCSKSRPLNDIVSQTTLOSLDITSRF 493
DB 242 FLGSNKRKRGALGNLKNSTYKKNNDNLSHFATITCSKSRPLNDIVSQTTLOSLDITSRF 301

OY 494 NSAIETALNRFIOKYDSYWRLLDDTSGK 521
DB 302 NSAIETALNRFIOKYDSYWRLLDDTSGK 329
RESULT 4
US-08-913-477-4
Sequence 4, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-4
Query Match 61.6%; Score 1641; DB 2; Length 329;
Best Local Similarity 99.7%; Pred. No. 1.8e-121;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 197 IRAYEQNPOHFEDEKRVVEQLTGHSSVLEELVQVYKDKNDISIKYDPRKDESEV 256
DB 5 IRAYEQNPOHFEDEKRVVEQLTGHSSVLEELVQVYKDKNDISIKYDPRKDESEV 64
OY 257 RVTDDIELKKILAYFLPEDTITLKGGHYDNOLONGIKRVKFELESSPNTOWELRAF 316
DB 65 RVTDDIELKKILAYFLPEDTITLKGGHYDNOLONGIKRVKFELESSPNTOWELRAF 124

QY 317 MRSSTADRIDDDILKVIYDSNHHGDAKSLKEELAEELAEIKIYSVIOAEINKHLSSS 376
|
Db 125 MHSSTADRIDDDILKVIYDSNHHGDAKSLKEELAEELAEIKIYSVIOAEINKHLSSS 184
QY 377 GTININDKSNLMDKNLYGTDEIEIFKASAEYKILEKMPOTTIOVGSSEKKIYSINDFLG 436
|
Db 185 GTININDKSNLMDKNLYGTDEIEIFKASAEYKILEKMPOTTIOVGSSEKKIYSINDFLG 244
QY 437 SENKRGALGNLKNYSYNKNDELHPATGSDKSRPLNDLVSOQKTTOLSDITSRFNSA 496
|
Db 245 SENKRGALGNLKNYSYNKNDELHPATGSDKSRPLNDLVSOQKTTOLSDITSRFNSA 304
QY 497 IEALNRFIOKYDSVMORLLDDTSGK 521
|
Db 305 IEALNRFIOKYDSVMORLLDDTSGK 329

RESULT 5

US-08-913-477-21
Sequence 21, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYTE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-21

Query Match 31.7%; Score 844; DB 2; Length 170;
Best Local Similarity 99.4%; Pred. No. 2,9e-59;
Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 MKRISVIAIAFGTATNADLTASTATATLTVEPARITLTVEGAPITMDNGNDIT 83
|
Db 1 MKRISVIAIAFGTATNADLTASTATATLTVEPARITLTVEGAPITMDNGNDIT 60
QY 84 ELLVGLTLGGYKGTGTSTSVNFTDAAGDPMYLFTFSODGNHOFITVYIGKDSRDFIS 143
|
Db 61 ELLVGLTLGGYKGTGTSTSVNFTDAAGDPMYLFTFSODGNHOFITVYIGKDSRDFIS 120
QY 144 PVRNGENLVGDVYVLTAGSQDFVRSIGSGKGLAAGYTTDAVTVTSNQ 193
|
Db 121 PVRNGENLVGDVYVLTAGSQDFVRSIGSGKGLAAGYTTDAVTVTSNQ 170

RESULT 6

US-08-913-477-11
Sequence 11, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYTE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-11


```

Db 117 KYFN--SISLNNEYTIINCENNSGKVSILNGEITITLDTQDEIK-----QRYVFK 166
QY 230 LVOLVR-----DKNIDISIKYDPRKDSVFANRVITDDIELLKKILAYFLPEDTILKGGH 284
Db 167 YSOMINISDYINRMIFVTITNNRLNLSKIYINGRLIDQ-----KPI-----207
QY 285 YDNQOLNGIKRVKFEFLSSPTQWEL-----RAFMAVHFSLTADRIDDLKVIYDSM 338
Db 208 --SNLGN-----IHASNNIMEKLDGCRDTHRYIWIKEFNFDEKLEKEIKDLQ 257
QY 339 NHGDARSKRLRELAETALKIYSYIOAEINKHLSSSGTINIHD--KSLNLDKNLYGY 396
Db 258 SNGS-----LKDFWGDLYQDKPYM-----LNLDPKRYVDVNNNGIRGY 299
QY 397 TDEIFKASAEYKILEKMPQTTIOVDGSEKKIYSINDFJGSENNKRGALGNKNSYSYNK 456
Db 300 -----MYLKGPRGSVMTNTIYLNSSLYR-GTKFLIKKVASGNK 336
QY 457 DN 458
Db 337 DN 338

RESULT 9
US-08-405-496A-26
Sequence 26, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
TITLE OF INVENTION: NEUROTOXIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear

```

```

MOLECULE TYPE: protein
US-08-405-496A-26

Query Match
Best Local Similarity 17.8%; Score 144; DB 2; Length 462;
Matches 86; Conservative 89; Mismatches 139; Indels 168; Gaps 22;

QY 1 MGHNNHHHHHHSSGHIDDDKMKKISSVIAIALGTLATANAADLTASTATATVLP 60
Db 1 MGHNNHHHHHHSSGHI--EGRMASMARL-----LSFTVEIKNIINT 42
QY 61 ARTLTLYKGAFTTINDNGNIDELLYGLTLAGYGTGTSVNTDAGDPMTLTFS 120
Db 43 SLNRYE-----SNHLIDLSRYASKINIG--SKVNF-----DPI-----75
QY 121 QDNHNOFTTKVIGKSDRDFDISPKVNGENLVDVYLATGSDPFVRSIGKGGKLANG 180
Db 76 -DKNQIQ-----LNFESSKIEVILK--NAIYNSMYENFSTFWIR-----IP 116
QY 181 KYTDAVTYVSNQEFMIRAYEONPQ-----HFIEDLEKRVAEOLTGHSVLEE 229
Db 117 KYFN--SISLNNEYTIINCENNSGKVSILNGEITITLDTQDEIK-----QRYVFK 166
QY 230 LVOLVR-----DKNIDISIKYDPRKDSVFANRVITDDIELLKKILAYFLPEDTILKGGH 284
Db 167 YSOMINISDYINRMIFVTITNNRLNLSKIYINGRLIDQ-----KPI-----207
QY 285 YDNQOLNGIKRVKFEFLSSPTQWEL-----RAFMAVHFSLTADRIDDLKVIYDSM 338
Db 208 --SNLGN-----IHASNNIMEKLDGCRDTHRYIWIKEFNFDEKLEKEIKDLQ 257
QY 339 NHGDARSKRLRELAETALKIYSYIOAEINKHLSSSGTINIHD--KSLNLDKNLYGY 396
Db 258 SNGS-----LKDFWGDLYQDKPYM-----LNLDPKRYVDVNNNGIRGY 299
QY 397 TDEIFKASAEYKILEKMPQTTIOVDGSEKKIYSINDFJGSENNKRGALGNKNSYSYNK 456
Db 300 -----MYLKGPRGSVMTNTIYLNSSLYR-GTKFLIKKVASGNK 336
QY 457 DN 458
Db 337 DN 338

RESULT 10
US-08-651-818A-19
Sequence 19, Application US/08651818A
Patent No. 5948889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
TITLE OF INVENTION: ANTIMICROBIALS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carrioll, Peter G.

```

REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-19

Query Match 5.0%; Score 132; DB 2; Length 53;
Best Local Similarity 73.3%; Pred. No. 0.00045;
Matches 22; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MGHNNHHHHSSGHI-----DDDDK 22
DB 1 MGHNNHHHHSSGHI-----DDDDK 30

RESULT 11
US-08-651-818A-23

Sequence 23, Application US/08651818A
Patent No. 5948889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
APPLICANT: Hale, Cynthia A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-23

Query Match 5.0%; Score 132; DB 2; Length 54;
Best Local Similarity 73.3%; Pred. No. 0.00047;
Matches 22; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 MGHNNHHHHSSGHI-----DDDDK 22
DB 1 MGHNNHHHHSSGHI-----DDDDK 22

DB 1 MGHNNHHHHSSGHI-----DDDDK 30

RESULT 12

US-08-188-228-42
Sequence 42, Application US/08188228
Patent No. 5597725
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,228
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/049,460
FILING DATE: 19 APR 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,643
FILING DATE: 17 APR 1992

ATTORNEY/AGENT INFORMATION:
NAME: No. 5597725and, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31340

TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300

TELEFAX: (312) 474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-188-228-42

Query Match 4.7%; Score 125.5; DB 1; Length 799;
Best Local Similarity 21.3%; Pred. No. 0.078;
Matches 107; Conservative 80; Mismatches 199; Indels 117; Gaps 25;

QY 19 DDDKMKKISSVIALALGTIATAAADLTASTATL--VEPARILT-----YKGA 71
DB 88 DDDPSKRIKIKYLSGDSGAGTIFQIN--DITGDIAHRIKRLDREKAEYLLTAQAVDMENK 145
QY 72 P-----ITMDNGNIDTELVGTLTGKVT-----GTSTSVNFTDADGDPYL 116
DB 146 PLEPSEFLIKVQDINDNAPELNGPY-----HATVPKSIIGTISTVNTATD-ADDPY- 199
QY 117 TFTSODGNNHOFITVKIKDSRDISPKV-----NGENLVGDVYLAATGSDFFVR 168
DB 200 -----GNSAKIVYSIL-EGOPYFSIEPEFTAIKITALPMRMRKAEYLVYQARD----- 248
QY 169 STGSGGLAGKYVDATVAVSNQEFMIRAVEQNPQHIEDLEVRVQGLHGSSYLE 228
DB 249 -MGHSGGLSG--TTTLTVLTVD-----VNDPPEFASLVHFSVPE-----DVLG 292
QY 229 ELVOLVKKNIDISIKYDPRDSEVFANRVITDIELLKLIAVFLPBDTLKGGHYDQ 288

Db 293 TAIGRVKANDODIGENAOSSD-----IIDGGTALFEITS-----DAQ 331
 QY 289 LONGIRKRVKEFLSSPTOMELRAFMAVH-----FSLTADRIDDILKIVYDSMNH----- 340
 Db 332 AODGVIRLRKPLDEFETKSKYTLKVEANAHIDPFRSGRPFKDTATVKIVEDADEPPVF 391
 QY 341 -----HGDAKSKLREELAEITAEK--IYSVIOAEINKH--LSSGTINIHKSN 387
 Db 392 SSPTYLLEVEHENA--ALNSVIGVYTAADPDITSSPIRESIDRHTDLEROFNINADDKIT 449
 QY 388 L--MDKNLYGTDEEIFRAS-AEYKILEKMPOT--TIQVDSSEKIVSINKDELGSENKR 441
 Db 450 LATPLDRELVSVMHNISIIATEIRNHSQISRPVPAIKVLVDNNAPEFASYEAFLECNCK 509
 QY 442 TGAIGLKNKSYSNKDNNELSHF 464
 Db 510 PGQV--IQTVSAMDKDDPKNGHF 530
 RESULT 13
 US-08-332-638-42
 ; Sequence 42, Application US/08332638
 ; Patent No. 5646250
 ; GENERAL INFORMATION:
 ; APPLICANT: Suzuki, Shintaro
 ; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; STREET: 6300 Sears Tower, 233 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/332,638
 ; FILING DATE: 01-NOV-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/872,643
 ; FILING DATE: 17 APR 1992
 ; APPLICATION NUMBER: US/08/049,460
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5646250and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 31340
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 474-6300
 ; TELEFAX: (312) 474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 799 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-332-638-42
 Query Match 4.7%; Score 125.5; DB 1; Length 799;
 Best Local Similarity 21.3%; Pred. No. 0.078;
 Matches 107; Conservative 80; Mismatches 199; Indels 117; Gaps 25;
 QY 19 DDDKHKKISSYALAFGLTATANAADLASTATATL--VEPARITLT-----YKEGA 71
 Db 88 DDDPSKKIKIYLSGGAGIITPOIN--DITGDIHAIKRLDREKAEYTLTAAQAVDMETNK 145

QY 72 P-----ITINDGNIDTELLVGTLLGKYT-----GTSTSVNFTDAAGDPMTL 116
 Db 146 PLPEPSEFIIKQYDINDNPEFLNGPY-----HATVPEMSILGTSYTNVATD-ADDPYV- 199
 QY 117 TETSQDGNHGFITTKYIGDSNDFDISPKV-----NGENLVGDVVLATGSODEFVR 168
 Db 200 -----GNSAKLYSIL-BQOPYFSTIEPTALIKTALPMDRPAKEEYLVYQAD----- 248
 QY 169 SIGSKGKLAAGKYTDVAVTVSNQEFMIRAVEONFQHIETDELEKRVQTLGHGSSYLE 228
 Db 249 -MGHSGGSLG--TTLVLTLD-----VNDNPKRAQSLXHSVPE-----DVYLG 292
 QY 229 ELVQLYKDKNIDISIKYDPRKDSVFANNVITDDIELLKILAVFLPEDTILKGYHDNQ 288
 Db 293 TAIGRVKANDODIGENAOSSD-----IIDGGTALFEITS-----DAQ 331
 QY 289 LONGIRKRVKEFLSSPTOMELRAFMAVH-----FSLTADRIDDILKIVYDSMNH----- 340
 Db 332 AODGVIRLRKPLDEFETKSKYTLKVEANAHIDPFRSGRPFKDTATVKIVEDADEPPVF 391
 QY 341 -----HGDAKSKLREELAEITAEK--IYSVIOAEINKH--LSSGTINIHKSN 387
 Db 392 SSPTYLLEVEHENA--ALNSVIGVYTAADPDITSSPIRESIDRHTDLEROFNINADDKIT 449
 QY 388 L--MDKNLYGTDEEIFRAS-AEYKILEKMPOT--TIQVDSSEKIVSINKDELGSENKR 441
 Db 450 LATPLDRELVSVMHNISIIATEIRNHSQISRPVPAIKVLVDNNAPEFASYEAFLECNCK 509
 QY 442 TGAIGLKNKSYSNKDNNELSHF 464
 Db 510 PGQV--IQTVSAMDKDDPKNGHF 530
 RESULT 14
 US-08-755-587-183
 ; Sequence 183, Application US/08755587
 ; Patent No. 6045997
 ; GENERAL INFORMATION:
 ; APPLICANT: Futreal, Phillip A
 ; APPLICANT: Wooster, Richard F
 ; APPLICANT: Ashworth, Alan
 ; APPLICANT: Stratton, Michael R
 ; TITLE OF INVENTION: Materials and methods relating to the
 ; TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
 ; NUMBER OF SEQUENCES: 222
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Bell Selzer Park & Gibson
 ; STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
 ; CITY: Raleigh
 ; STATE: NC
 ; COUNTRY: USA
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/755,587
 ; FILING DATE: 25-NOV-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9523959.6
 ; FILING DATE: 23-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9525555.0
 ; FILING DATE: 14-DEC-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9617961.9
 ; FILING DATE: 28-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kenneth D Sibley
 ; REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5405-135
 INFORMATION FOR SEQ ID NO: 183:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1683 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-755-587-183

Query Match 4.7%; Score 124; DB 3; Length 1683;
 Best Local Similarity 18.4%; Pred. No. 0.3; Mismatches 218; Gaps 34;
 Matches 125; Conservative 120;

13 SSCGI-DGDDKMKKISSVI-----AIALEGTATNADLSTATATLV 58
 575 ASGLTDENEVERGFSALGTKNVSSALQKVKLF-----SDIENISEIAEV 626
 59 EPARITLT--YKGAFTINDNGNID-----TELLVGT--LTIG-----GYKT 97
 627 DPISLSSSKYHDSVASMKIENQXDKSEKNNKCOLLNNIMMTGIFVEENTENYKR 686
 98 GTSTSTVNTDAAGDPYVLTFTSQDGNHQTFTVIGKDRDF-----DISKRVNGEN 150
 687 NTEHEDNKRYTGASRNSXNLE--XSDGSDSSKNDVTYIHKDETLPFIQHNKICLKLGOF 745
 151 LVGDDVVLATGODEFYSIGSKGKLAAGKYDPAVTYVTSNOEFMIRAVE----- 201
 746 MKEGNOIKRGLSD--LTCLYMAKEETCHGNTSKREQLTAKTQNKIKDPTPISTQTA 803
 202 --ONPQHFIEDL-----EKVAREQLTGHSSVLELVQVYKDKNIDIS----- 242
 804 SGNKIRVSKESLNAVNEFPQKXTEELNNSDSLNSSELLSGINKNMKDISHXETDI 863
 243 IKYDPRDSEY-FANRYIT-----DIELLKRTIAFLPBDITLKGHNDQNGIKR 295
 864 VKNKILKESPVGTGNOLVTLQORECEIEKIK-----EPTLL-GFHNASGRK--VKI 912
 296 VKFELESSPTQWELRAFMAMVHFLPADRIDDLKVIYDSNMHGHGASKLE----- 350
 913 AKESLDKVKML-----FDETEQGVSE-----ITSESHQAKITAKDEACKDG 954
 351 -----ELAEFLAEIKIYVTOAEINKHLSSSGTI-----NIHDSINLMDKN----- 392
 955 LELACEVEITAPKCEMONGSLNDKYSKEITVLPQLLSNDLYRQENLKTNSLSLK 1014
 393 --LGYTDEELFKR-----SAEYKILEKMP-----QTTIQ----- 420
 1015 VKHVENVEKETAKSPCTCYTNSQSYVIEGSLAFYTGHSRKTVSSEASLLEAKKWLREG 1074
 421 -VDGSEKKIYSIK-----DFLG-----SENKRTGALGN--LK 449
 1075 IFDQDPRINTAKYCEKEHEDYVGNALYENSNSITENDKNHLSERKODSTIYNSMS 1134
 450 NSYSYNDK-----NNELSHFATTCSDK--RP-LNDLYSQKTTQDLSRPSNAIEA 499
 1135 NSYSYSHSDFCHSDDEVYNDGSLSKNKIDXSIEFLVKNVEDQKNISSEV-----SAVKE 1190
 500 LNRFIQKIDSY-MORLDDIS 519
 1191 ANTPOTVNDICVEKLVNNS 1211

RESULT 15
 US-08-480-604A-24
 Sequence 24, Application US/08480604A
 Patent No. 5736139
 GENERAL INFORMATION:
 APPLICANT: KINK, JOHN A.
 APPLICANT: THALLEY, BRUCE S.
 APPLICANT: PADHYE, NISHA V.
 APPLICANT: FIRCA, JOSEPH R.
 APPLICANT: STAFFORD, DOUGLAS C.
 TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND

TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MEDLEN & CARROLL, LLP
 STREET: 220 MONTGOMERY STREET, SUITE 2200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/480,604A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/422,711
 FILING DATE: 14-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/405,496
 FILING DATE: 16-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/329,154
 FILING DATE: 25-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/429,791
 FILING DATE: 31-OCT-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: INGOLTA, DIANE E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: OPHD-01763
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 23 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-604A-24

Query Match 4.6%; Score 122; DB 1; Length 23;
 Best Local Similarity 79.2%; Pred. No. 0.00081;
 Matches 19; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
 1 MGHNNHHHHHSSGHTDDDKRM 24
 1 MGHNNHHHHHSSGHT--EGRHM 22

Search completed: August 22, 2000, 16:47:15
 Job time: 1092 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 16:32:34 ; Search time 38.09 Seconds

(without alignments)
846,468 Million cell updates/sec

Title: US-08-699-716a-2

Sequence: 2666 1 MGHNNHHHHHHSSGHIDD.....RFIOKYDSVMORLDDITSGK 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 178050 segs, 61884766 residues

Total number of hits satisfying chosen parameters: 178050

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1646	61.7	326	2 T43594	yop secretion and
2	1645	61.7	326	2 B33601	lcrv protein - yer
3	1590	59.6	326	2 B37314	regulatory protein
4	850	31.9	170	2 S13008	capsular antigen F
5	164.5	6.2	1252	2 B42771	reticulocyte-bind1
6	150.5	5.6	1116	2 S57382	hypothetical prote
7	149	5.6	2166	2 G70163	hypothetical prote
8	147	5.5	785	2 T50180	cadherin-7 - chick
9	145	5.4	964	2 S57379	MSH2 protein - yea
10	144	5.4	1115	2 T41342	probable coiled-co
11	139.5	5.2	1957	2 T38077	hypothetical colle
12	138.5	5.2	644	2 T46211	hypothetical prote
13	137	5.1	1979	2 C71622	hypothetical prote
14	136	5.1	1312	1 BMBYRL	RAD50 protein - ye
15	136	5.1	2269	2 T28677	ripty protein - ye
16	135	5.1	1173	2 T25539	hypothetical prote
17	135	5.1	1510	2 T16927	myosin heavy chain
18	135	5.1	1938	1 MMKWL	myosin heavy chain
19	135	5.1	1938	2 T21193	hypothetical prote
20	134.5	5.0	1467	2 PC1253	tyb protein - yeas
21	134	5.0	1939	2 T18372	repeat organellar
22	133.5	5.0	2401	2 T28676	ripty protein -
23	133	5.0	1039	2 S62509	probable vesicular
24	132	5.0	168	2 T58315	w11 - human
25	132	5.0	2606	2 T03159	large tegument pro
26	131.5	4.9	547	2 G64851	flagellar hook-ass
27	131	4.9	1803	2 S56894	tyb protein - yeas
28	130.5	4.9	1427	2 S22695	reslin - human
29	130	4.9	551	2 T25197	hypothetical prote

ALIGNMENTS

RESULT 1
T43594

yop secretion and targeting control protein - Yersinia pestis plasmid pCD1

C:Species: Yersinia pestis
C:Date: 21-Jan-2000 #sequence-revision 21-Jan-2000 #text-change 04-Mar-2000

C:Accession: T43594; T42889
R:Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Bruba

J. Bacteriol. 180, 5192-5202, 1998
A:Title: Structural organization of virulence-associated plasmids of Yersinia pestis.

A:Reference number: 222578; M0ID:98422474
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-326 <HUP>
A:Cross-references: EMBL:AF053946; PIDN:AC62574.1

A:Experimental source: strain KIM
R:Periy, R.D.; Straley, S.C.; Fetherston, J.D.; Rose, D.J.; Gregor, J.; Blattner, F.R

Infect. Immun. 66, 4611-4623, 1998
A:Title: DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersini

A:Reference number: 222573; M0ID:98427122
A:Accession: T42889

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-326 <PER>

A:Cross-references: EMBL:AF074612; PIDN:AC69799.1
A:Experimental source: strain KIM5

C:Genetics:
A:Gene: lcrv
A:Genome: plasmid pCD1

Query Match 61.7%; Score 1646; DB 2; Length 326;
Best Local Similarly 99.7%; Pred. No. 1.3e-79;
Matches 325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	196	MIRAEYONPOHFEDLEKRYEOLTGHSVLEELVQVYDKNDISIKDPKRDSEYFA	255
DB	1	MIRAEYONPOHFEDLEKRYEOLTGHSVLEELVQVYDKNDISIKDPKRDSEYFA	60
QY	256	NRVTIDIELKRLILAYFLPEDITLKGHYDNOQNGIKRYKEFLSSPNTOWELRAFWA	315
DB	61	NRVTIDIELKRLILAYFLPEDITLKGHYDNOQNGIKRYKEFLSSPNTOWELRAFWA	120
QY	316	VMHSLADRIDDIILKYVDSNMHGDASKLKEELAEITAEIKTISVQAEINKHLSS	375
DB	121	VMHSLADRIDDIILKYVDSNMHGDASKLKEELAEITAEIKTISVQAEINKHLSS	180
QY	376	SGTINIHDSKINLMDKNLYGTDEIFKASAEYKILEKMPOTTIOVGSSEKIYSINDFL	435
DB	181	SGTINIHDSKINLMDKNLYGTDEIFKASAEYKILEKMPOTTIOVGSSEKIYSINDFL	240
QY	436	GSNKRKRGALGNLKNYSYKNNKNNELSHFATCTCSDKSRPLNDLVSOQTOLSDITSRFS	495

conserved hypothet
translation elonga
translation elonga
spindle pole body
kinesin-related pr
tyb protein - yeas
myosin heavy chain
hypothetical prote
inner layer protei
tyb protein - yeas
conserved hypothet
hypothetical prote
myosin-like protei
ribosomal protein
hypothetical prote
flagellar p-ring p

```

Db      241 GSENRKTGALGNLKNKSYNKNLSEHFAITCSDKSRPLNDLVSOQTQLSDITSRFS 300
      496 ATEALNRFIOKYDSVMORLLDPTSGK 521
      301 ATEALNRFIOKYDSVMORLLDPTSGK 326

RESULT 2
lcry protein - Yersinia pestis
C:Species: Yersinia pestis
C>Date: 17-Jan-1990 #sequence_revision 17-Jan-1990 #text_change 23-Mar-1993
C:Accession: B33601
R:Price, S.B.; Leung, K.Y.; Barve, S.S.; Straley, S.C.
J. Bacteriol. 171, 5646-5653, 1989
A:Title: Molecular analysis of lcryVH, the V antigen operon of Yersinia pestis.
A:Reference number: A33601; MUID:90008806
A:Accession: B33601
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-326 <PR1>
A:Cross-references: GB:M26405

```

```

Query Match          61.7%: Score 1645; DB 2; Length 326;
Best Local Similarity 99.4%: Pred. No. 1.5e-79;
Matches 324; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      196 MIRAYQNPQHFEDELEKRVLEQTLGHGSSVLEELVQVKNKIDISIKYPRKDESEVA 255
      1 MIRAYQNPQHFEDELEKRVLEQTLGHGSSVLEELVQVKNKIDISIKYPRKDESEVA 60
QY      256 NRYITDDIELKKILAYFLPEPTILKGGHYDNLQNGIKRVKEFLLESSPNTQWELRAFWA 315
      61 NRYITDDIELKKILAYFLPEPTILKGGHYDNLQNGIKRVKEFLLESSPNTQWELRAFWA 120
QY      316 VWFSLTADRIDDDILKTVYDSNNHNGDARSKLRELAETLTKYVYQAEINKHLSS 375
      121 VWFSLTADRIDDDILKTVYDSNNHNGDARSKLRELAETLTKYVYQAEINKHLSS 180
QY      376 SGTINIHDSINIMDKNLGYTDEEIFKASAEYKILEKMPOTTIOVDSSEKIVSIKDFL 435
      181 SGTINIHDSINIMDKNLGYTDEEIFKASAEYKILEKMPOTTIOVDSSEKIVSIKDFL 240
QY      436 GSENRKTGALGNLKNKSYNKNLSEHFAITCSDKSRPLNDLVSOQTQLSDITSRFS 495
      241 GSENRKTGALGNLKNKSYNKNLSEHFAITCSDKSRPLNDLVSOQTQLSDITSRFS 300
QY      496 ATEALNRFIOKYDSVMORLLDPTSGK 521
      301 ATEALNRFIOKYDSVMORLLDPTSGK 326

```

```

RESULT 3
B37314
regulatory protein lcryV - Yersinia pseudotuberculosis
C:Species: Yersinia pseudotuberculosis
C>Date: 30-Jan-1993 #sequence_revision 30-Jan-1993 #text_change 08-Oct-1999
C:Accession: B37314
R:Bergman, T.; Hakansson, S.; Forsberg, A.; Norlander, L.; Macellaro, A.; Baeckman, A.;
J. Bacteriol. 173, 1607-1616, 1991
A:Title: Analysis of the V antigen lcryVH-yopBD operon of Yersinia pseudotuberculosis:
A:Reference number: A37314; MUID:91154114
A:Accession: B37314
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-326 <BER>
A:Cross-references: GB:M57893; NID:g155456; PIDN:AAA27645.1; PID:g155458

```

```

Query Match          59.6%: Score 1590; DB 2; Length 326;
Best Local Similarity 96.6%: Pred. No. 1.1e-76;

```

```

Matches 315; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY      196 MIRAYQNPQHFEDELEKRVLEQTLGHGSSVLEELVQVKNKIDISIKYPRKDESEVA 255
      1 MIRAYQNPQHFEDELEKRVLEQTLGHGSSVLEELVQVKNKIDISIKYPRKDESEVA 60
QY      256 NRYITDDIELKKILAYFLPEPTILKGGHYDNLQNGIKRVKEFLLESSPNTQWELRAFWA 315
      61 NRYITDDIELKKILAYFLPEPTILKGGHYDNLQNGIKRVKEFLLESSPNTQWELRAFWA 120
QY      316 VWFSLTADRIDDDILKTVYDSNNHNGDARSKLRELAETLTKYVYQAEINKHLSS 375
      121 VWFSLTADRIDDDILKTVYDSNNHNGDARSKLRELAETLTKYVYQAEINKHLSS 180
QY      376 SGTINIHDSINIMDKNLGYTDEEIFKASAEYKILEKMPOTTIOVDSSEKIVSIKDFL 435
      181 SGTINIHDSINIMDKNLGYTDEEIFKASAEYKILEKMPOTTIOVDSSEKIVSIKDFL 240
QY      436 GSENRKTGALGNLKNKSYNKNLSEHFAITCSDKSRPLNDLVSOQTQLSDITSRFS 495
      241 GSENRKTGALGNLKNKSYNKNLSEHFAITCSDKSRPLNDLVSOQTQLSDITSRFS 300
QY      496 ATEALNRFIOKYDSVMORLLDPTSGK 521
      301 ATEALNRFIOKYDSVMORLLDPTSGK 326

```

```

RESULT 4
S13008
capsular antigen F1 precursor - Yersinia pestis plasmid pMT1

```

```

C:Species: Yersinia pestis
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 04-Mar-2000
C:Accession: S13008; T14702; T15015; S23725
R:Galayov, E.E.; Smirnov, O.Y.; Karlishhev, A.V.; Volkovoy, K.I.; Denesynuk, A.I.; Nazim
FBS Lett. 277, 230-232, 1990
A:Title: Nucleotide sequence of the Yersinia pestis gene encoding F1 antigen and the
A:Reference number: S13008; MUID:91099503
A:Accession: S13008
A:Molecule type: DNA
A:Residues: 1-170 <GAL>
A:Cross-references: EXBL:X61996; NID:g48620; PIDN:CAA43966.1; PID:g48621
R:Hu, P.; Elliott, J.; McCready, P.; Showronski, E.; Garnea, J.; Kobayashi, A.; Carra
submitted to the EMBL Data Library, March 1998
A:Description: Structural organization of virulence determinants in three Yersinia pe
A:Reference number: T14702
A:Accession: T14702
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-170 <HUP>
A:Cross-references: EXBL:AF053947; NID:g2996286; PID:g2996338; PIDN:AA013218.1
R:Lindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.
Infect. Immun. 66, 5731-5742, 1998
A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIMS plas
A:Reference number: Z18268; MUID:99043898
A:Accession: T15015
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-170 <LIN>
A:Cross-references: EXBL:AF074611; NID:g3883003; PID:g3883098; PIDN:AA082758.1
C:Genetics:
A:Gene: cafI
A:Genome: plasmid pMT1
C:Superfamily: Yersinia pestis plasmid pMT1 capsular antigen F1
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-170/Product: capsular antigen F1 #status predicted <MAT>

```

```

Query Match          31.9%: Score 850; DB 2; Length 170;
Best Local Similarity 100.0%: Pred. No. 3e-38;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      24 MKKISSVIAIAALFGTIAANAADLTASTTATATLVEPRATITLYTEGAPITIMONGNIDT 83
      1 MKKISSVIAIAALFGTIAANAADLTASTTATATLVEPRATITLYTEGAPITIMONGNIDT 83

```

Db 1 MKKISSVIAIAFGTATATANAADLTASTATATLVEPARITLLYKGGAPITIMDNGNIDT 60
 Qy 84 ELLVGLTLLGGYKTTGTTSTSVNFTDAAGDPMYLLFTSODGNHQTFTKYIGKDSRFDIS 143
 Db 61 ELLVGLTLLGGYKTTGTTSTSVNFTDAAGDPMYLLFTSODGNHQTFTKYIGKDSRFDIS 120
 Qy 144 PKVNGENLVGDVYVLTATGSDOFFVRSIGSGKGLAAGKTYTDATVTVSNQ 193
 Db 121 PKVNGENLVGDVYVLTATGSDOFFVRSIGSGKGLAAGKTYTDATVTVSNQ 170

RESULT 5
 B42771
 reticulocyte-binding protein 2 - Plasmodium vivax (fragment)
 C:Species: Plasmodium vivax
 C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Sep-1997
 C:Accession: B42771
 R:Galinski, M.R.; Medina, C.C.; Ingravallo, P.; Barnwell, J.W.
 Cell 69, 1213-1226, 1992
 A:Title: A reticulocyte-binding protein complex of Plasmodium vivax merozoites.
 A:Reference number: A42771; MUID:92315338
 A:Accession: B42771
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1252 <GAL>
 A:Cross-references: GB:M88098; NID:q160627; PID:q160628
 A:Experimental source: strain Belem, merozoites
 C:Genetics:
 A:Gene: RBP2

Query Match 6.2%; Score 164.5; DB 2; Length 1252;
 Best Local Similarity 20.3%; Pred. No. 0.41;
 Matches 131; Conservative 104; Mismatches 221; Indels 189; Gaps 33;

Qy 8 HHHHSSGHIDDD-----KHKKISSVIAIAFGTATANAADL----- 47
 Db 402 HHHHSSGHIDDD-----KHKKISSVIAIAFGTATANAADL----- 47
 Qy 48 -TASTATATL-----VEPARI-----TLVYKGGAPITIM 76
 Db 460 ATSENTAEKLEKVDQSVYVYLNQITTEKRLVTEKRLGIDSTINIGA--LKE 517
 Qy 77 DNGNIDTELLVGLTLLG---GYKTTGTTSTSVNFTDAAGDPMYLLFTSODGNHQTFTKYI 133
 Db 518 SKGNVYIGLEKLEIGKNNKLVDTIKKINSTGVNFSLFNNF----- 562
 Qy 134 GKDSRFDISPKVNG--ENVLGDVYVLTATGSDOFFVRSIGSGKGLAAGKTYTDATVTVSN 192
 Db 563 --DLNQYDENKKNINDYENKKEIYNFEFGLNKISENL--RNASENTSDYNSAKTLRLA 618
 Qy 193 OEEMIRAY--EONPOHFIELEK-----RYEQUL-----GHSS--S 225
 Db 619 OKKAVLNKKEEANKYLRVKKVYVESFRLFNKESLIDKINEMIKKQULVNEGHWVKO 678
 Qy 226 VLEELVQVYKKNIDISIKYDPRKDEVEFANRYITDDIELKKILAVFLPE--DTLLKGG 283
 Db 679 LYENIKELVDENLSDILKQATGKNEI-----OKITSTLKNKAKITIL--G 723
 Qy 284 HYDNOQLN-GIKRYAEF-----LESSPPTOMELR--AFMAVHFSLADRID-- 327
 Db 724 HDVTSKAYVGIKITPELALTELLGDAKLKTAQELKFESEKNNVLETEMKKNTELDVHK 783
 Qy 328 --DILKVIYDSMNNHG--DARSKLBELELAELKIVYVIAEINKHLSGGTINIH 393
 Db 784 NIDDAKVALEILAHDEIDTKQDSKLEMGNOYLKAVL--INQYNNKISSISKE 840
 Qy 384 KSNLMDKNLYGYTDEIFKASAEK-----ILKAPQTTIO-----VDGSEK 426
 Db 841 EAVSVKIGNV-SKKHSELKITCSKSYDNIILAK--QTELDNLNRSFTQEKTNNSDS 897
 Qy 427 KIYISIK-DLFGSENKRTGALGNLK--NSYSYKDNNELSHFATITCSDKRPLNDLVSO 482

Db 898 KLEKIRTDESLEKN-----ALKTLEGEVNAALKASDNHE-----HYQSKSEPVNPALSEI 947
 Qy 483 TTOLDITSRFNSALEAL-----NRFIOKYDSVMORLDDT 518
 Db 948 EKEETDIDS-LMTALDELKAKRTECVSRKLIKIDVTNIEISDDT 991

RESULT 6
 S57382
 hypothetical protein YOL087c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein 00944
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 12-Dec-1997
 C:Accession: S57382; S66781; S50418
 R:Zumstein, E.; Pearson, B.M.; Kaloogeropoulos, A.; Schweizer, M.
 Yeast 11, 975-986, 1995
 A:Title: A 29.425 kb segment on the left arm of yeast chromosome XV contains more than
 A:Reference number: S57374; MUID:96021609
 A:Accession: S57382
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-1116 <ZUM>
 A:Cross-references: EMBL:X83121; NID:q600461; PID:q600471
 R:Zumstein, E.; Pearson, B.M.; Kaloogeropoulos, A.; Schweizer, M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66775
 A:Accession: S66781
 A:Molecule type: DNA
 A:Residues: 1-1116 <ZUM>
 A:Cross-references: EMBL:Z74829; NID:q1419927; PID:q251885; PID:q1419928; MIPS:YOL087
 A:Experimental source: strain S288C
 C:Genetics:
 A:Map position: 15L

Query Match 5.6%; Score 150.5; DB 2; Length 1116;
 Best Local Similarity 18.3%; Pred. No. 1.9;
 Matches 94; Conservative 96; Mismatches 193; Indels 131; Gaps 20;

Qy 3 HHHHHHHHSSGHIDDDKHKKISSVIAIAL-----FGTATANAADLTASTAT 53
 Db 365 HHHHHHHHEBONISTTAKKYGGLIDIALLPNEKLLFECDDNSNVLDTLNHNS 424
 Qy 54 TATVEPARITLLYKGGAPITIMDNGNIDTELLVGLTLLGKYKTTGTTSTSVNFTDAAGP 113
 Db 425 ----VNEGFALTRSS-----LTNRHVVITENTKGMOMWDYSCLELTFPSSBSGSPD 476
 Qy 114 MYLFTSODGNHQT-----TKVIGKDSRFDISPKVNGENLVGD 155
 Db 477 IVKRYTSKEILSHWCIVSVYKGLFVKINPKFLKTEVYSALKDQY--VNNIEINSDE 533
 Qy 156 -----VLTATGSDOFFVRSIGSGKGLAAGKTYTDATVTVSNQOEEMIRAYEONPOHFIE 209
 Db 534 RYNLGKIVINSLENERI--SYEQKDKLKRKIFSLKRDNLNSLTDLDGYSESSEKNNK 591
 Qy 210 DLEKRVEDLT-----GHSS-----VLEE--LVOLVKKNIDISIK 244
 Db 592 DKRKRTFKISSTLSIGNNNSGTPPNSAPATPAVMAETVLEEQPLQASDKAID---- 647
 Qy 245 YDPRKDSVFANRVITDDIELKKILAYFLPEDTILKGYHYNQONGI--KRVKEFLES 302
 Db 648 -----DSELEVQPLPASKP-----YFRIOSSGSLSRKFRFRST 683
 Qy 303 SP-----NTOMELRPAFVAMHESLTADRIDDLKVIYDSMNNHGDAAS-----K 347
 Db 684 SGRTATGLTWPEPKGILDTDRVIN--DASAPQANTQOOSKADAPESKLMNHPK 739
 Qy 348 LREELAELETAEE-----IKIVSVIAEINKHLSGGTINIHDS-----ILMDKNLYGYD 398
 Db 740 LEOKLSAISQDLPNSNTNKNLRSNSRANSTSLIEGKAKKPPMPDLQIOESYKO 799
 Qy 399 EEIFKASAEYKILEKMPQTTIOVDSGEKKIVSIK 432

Db 800 QYNTSSSLKY-LTKRLPYTKI-IRASSCPILRVK 831

RESULT 7

hypothetical protein BR0512 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: G70163

A:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vuglt; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-2166 <L>

A:Cross-references: GB:AE001153; GB:AE000783; NID:92688419; PIDN:AA066876.1; PID:9268842 A:Experimental source: strain B31

Query Match 5.6%; Score 149; DB 2; Length 2166;

Best Local Similarity 21.6%; Pred. No. 5.6;

Matches 111; Conservative 85; Mismatches 154; Indels 164; Gaps 31;

75 IMDNGNIDTELLVGLTGLGKGTGTTSTVNFDAAGDPMYL-----TPTS----- 120

Db 890 LSSGKVDLDIDSEV-----TKIKELKFSLESLEYLEKIDERNQAGIYSDLL 942

QY 121 QDGNHQTIVKIKDSRDP--ISPKV-----NGENLVGD-----DVLATGSG 163

Db 943 QDINM-----FKETRELEENISKKFAVLNSEEVEKVDLQDRIDIASFOANI 996

QY 164 DEFVRSIGSKG--GKLAGKYTDVTVTSNOEFMIRAYQONQHIEDLEK---VAVE 217

Db 997 DITDLSLVKFNINKEIKNGKINVEITISYRGYSENISKLENEIMHETENLSRRLTDRI 1056

QY 218 QLTGHGSSVLELVQVQKNDISIKYPRKDEVFANRY--ITDIE-----LTKIL 270

Db 1057 SL-----SKGDENELQKKE--SEFVS-KYQVEK-----PELAKYDLDLQDEAKINLVKIE 1106

QY 271 AYFLP-----EDTLKQ--GHYNOLQNGIKRVEKFLSSPTQWELRAF 314

Db 1107 QYYSRLEEAIDYRTIDNDIMQAKEREGETNELKNIESSEFL-----NDLYKER--- 1159

QY 315 AVMHFSLADRIDDDILKVIYDSMHHGDARSKBELAELTALAKIYSVIAQAINKHL 374

Db 1160 -----FKLIESNFEERYSTFLESEG-----ATSKIRDE-----LYKTL----- 1193

QY 375 SSGTINIHDSKINIMDKRNLKYGTDEIFKASAEKILKEMPOITIOVDSSEKIV-SIKD 433

Db 1194 TSNDEINQIK-ISEMDN-----PEIIFQKRIKILF--EKELQDKIKD 1234

QY 434 FLGSEKRTGAL-----GNLKSYSYNKDNELSHFATTCSDKSRPL--NDLVQSKYTQ 486

Db 1235 CYGRINSQFGKIKAGEVENIKN-----HF-DVCIKRVNLTLDIDIVKYE----- 1277

QY 487 SDIRSNSAIEALNRFIOKDKSVQRLDLDTS 520

Db 1278 -----NEIHKRIDSLKSTSTFDSIEKMLNDKVS 1307

RESULT 8

150180

cadherin-7 - chicken

C:Species: Gallus gallus (chicken)

C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000

C:Accession: I50180

R:Nakagawa, S.; Takeichi, M. Development 121, 1321-1332, 1995

A:Title: Neural crest cell-cell adhesion controlled by sequential and subpopulation-s

A:Reference number: I50178; MUID:95309115

A:Accession: I50180

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-785 <NA>

A:Cross-references: GB:D42150; NID:9868000; PIDN:BA07721.1; PID:9868001

C:Superfamily: cadherin; cadherin repeat homology

F:156-262/Domain: cadherin repeat homology <CDH>

Query Match 5.5%; Score 147; DB 2; Length 785;

Best Local Similarity 23.0%; Pred. No. 1.8;

Matches 115; Conservative 71; Mismatches 201; Indels 114; Gaps 26;

QY 12 HSSGHIDDDKHKMKI--SVYIALFETIATANAADLTAS-----TTATATVEPARIT 64

Db 72 HSD--VDKGGSGIKYILSGGASIF--IDENTGDIHAKRKDRDREGAYTTLRQAHDR 127

QY 65 LTYEGAP-----ITMDNGNIDTELLVGLTGLGKX--TGTTSTVNFDAAGDPMYL 116

Db 128 LTNKPVPESEFVKIIDIINDEPKFLDGYTAGVPSPVGSVQVATD--ADPTY- 185

QY 117 TPTSDGNHQTIVKIKDSRDPDISPKV-----NGENLVGDVYVATGSDFFVR 168

Db 186 -----GNSARVYSLI--QGQPFYSEPKTGIIKTALPNDRKADQYLLVIAQDM-- 235

QY 169 SIGSGGKLAGKYTDVTVTSNOEFMIRAYQONQHIEDLEKVEQLTGHGSSVLE 228

Db 236 -VGONGG--LSG--TTSVYTLID-----VNDNPRFRRSYQVYVPE-----SLPLA 278

QY 229 ELVOLVQKNDISIKYPRKDEVFANRYITDIELKILAYFLPEDTILKGHDNQ 288

Db 279 SVVARIKADADVG---PNAEMEY---KIVDGGGLGVFKI-----SVDKD 317

QY 289 LQNGIKRVEKFLSSPTQWELRAFAMVH-----FSLTARIDDDILKVIYDSMHHGDA 344

Db 318 TQBITITQKELDEAKTSTLRIEAMNHVDRPFLSLGFSMTYTKIIVEDV----- 371

QY 345 RSKRLAEALTEALIKIYSVIOAEINKHLSSGTINIHDSIN-----LMDNKLGYTD 398

Db 372 -----DEPPYFVS--RISNVVSEAKVGTITIGVAHADDAHSNPRYSIDRN-----TD 420

QY 399 -EELFKASAEYKILKEMPOITIOVDSSEKIVSIKDFLGSSEKRTGALGNLKSYSYNKD 457

Db 421 LERFEINIDANSVY-----TTAKSLDRETNVAVNITVLAMESQNPQIGGYVAITLID 475

QY 458 NNELSHFA-----TTCSDKSRP 474

Db 476 NDNAPEFAMEYETTVCEAP 496

RESULT 9

MSH2 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein O0935; protein YOL090W

C:Species: Saccharomyces cerevisiae

C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Sep-1999

C:Accession: S57379; S50415; S50782; S56784; S27433

R:Zumstede, E.; Pearson, B.M.; Kaloogeropoulos, A.; Schweitzer, M. Yeast 11, 975-986, 1995

A:Title: A 29.425 kb segment on the left arm of yeast chromosome XV contains more tha

A:Reference number: S57374; MUID:96021609

A:Accession: S57379

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-964 <200>

A:Cross-references: EMBL:883121; NID:9600461; PIDN:CA58189.1; PID:9600468

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994

R:Zumstede, E.; Pearson, B.M.; Kaloogeropoulos, A.; Schweitzer, M. submitted to the EMBL Data Library, December 1994

A:Reference number: S50410

A:Accession: S50415

hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T38077

A:Conor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, April 1996

A:Reference number: 221767

A:Accession: T38077

A:Status: Preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-1957 <CON>

A:Cross-references: EMBL:270690; PIDN:CAA94624.1; GSPDB:GN00066; SPDB:SPAC1F3.06C

A:Experimental source: strain 972h; cosmid c1f3

C:Genetics:

A:gene: SPDB:SPAC1F3.06C

A:Map position: 1

Query Match

Best Local Similarity 21.8%; Score 139.5; DB 2; Length 1957;

Matches 103; Conservative 74; Mismatches 161; Indels 135; Gaps 21;

QY 145 KVGNGENLV-----GDDV-----VLATGSDFFVRSIGSGKGLAAGKYDAVT 187

DB 1021 KNGENIASIQTETEKRAENDLOSLVSSSEYENLL--ISSQTNKSLDK----- 1072

QY 188 VYVSNDFMIRAYEONFOHIE--DLEKRVESQITGHSVLEELVOLVDRKNIDISIKY 245

DB 1073 ---TNQ---LKYTEKNVOKLDEKQORNELEELTSKYGKLGEBNAQ--IKDEL--LALRK 1123

QY 246 DPRDSVFANRYTDDIELLKLIAFLPED-----TILKGHDND-----LONGI 293

DB 1124 KSKQKQDLCAFW--DULKEKSDALEQITNEKNELVLSQNSNNALVEERSDILNRL 1181

QY 294 KRYKEFLSSPNTQWELRAFWAVHFSITADRIDDIILKY-----IYDSMNHG 342

DB 1182 SDKKSLSDSDNIVSYRSDLVRYNDELDTLKKDKDSLSTQYSVCODRDDLDLSLGC 1241

QY 343 DANKLEELAEIT-----AEIKIYVIAQAEINRHLSSSGT 378

DB 1242 ESNKRYAVSLRELCTKSEIDVPVSEILDNFVFNAGNSELSRLTVLSLE--NYLDAFNQ 1299

QY 379 INHDSINLMQNLVYTDSEIFKASAEKILEKMPQ-----TIQVDSSEKIV--SIKD 433

DB 1300 VNRKHELD---NRLTTDAEFTKVVAD---LEKIQHEHDDWLIQGDLEKALKDSEKN 1352

QY 434 FLGSENRRTGALNKL-----NSYSYKDNNEHSHFATTCSDSRP 474

DB 1353 FLKREAMTENHSLSEKRETKKEIAELSSRLLEDNOLANKLKNQDH----- 1401

QY 475 LNDLVSKTQQLSDITSRFNSAIEALNRFIOKIDSV-----NORLDDTSGK 521

DB 1402 LNEIRLEKEDVLKEKELIISLESISLNOROKESSLIDAKNEHEHMDTSRK 1454

RESULT 12

T46211

hypothetical protein TBP19.180 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46211

R:Choi, N.; Robert, C.; Bottler, P.; Winkler, P.; Cattolico, L.; Artiguenave, F.; Sa

submitted to the protein Sequence Database, December 1999

A:Reference number: 223008

A:Accession: T46211

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-644 <CHO>

A:Cross-references: EMBL:AL133315

C:Genetics:

A:Experimental source: cultivar Columbia; BAC clone TBP19

A:Map position: 3

A:introns: 133/1; 373/3; 403/3; 496/3; 566/3

A>Note: TBP19.180

Query Match

Best Local Similarity 19.9%; Score 138.5; DB 2; Length 644;

Matches 113; Conservative 91; Mismatches 211; Indels 153; Gaps 27;

QY 10 HHSSGHIDDDKHKRISSVIALEGTATANAADLTAATATATLVEPARITLYRE 69

DB 67 HAGVGSNSDKRSAREKASHLALVKY-----LODLADSSASEPSS--SKRKN 114

QY 70 GAITTINDN-----NIDELVGLTTLGCGYTGITSTYV-----NTDA 110

DB 115 GNPIDQCDHDEKLVPWKGIIVNIPF-----TAODGRAGEGSKLRDEYILRGFNPT 169

QY 111 GDPMY-----LFTSODGN-----NHQFTTVYIGKSRDPSIPKV----- 146

DB 170 VRELMTYLGHSATVIEPKNDNGLNGLLFDKAYIVDGHGKDWLKKQPKLGLYGTIA 229

QY 147 -----NGENLVGDDVVLATGSDFFVRSIGSGKGLAAGKYDAVTAVTSNOEFMIRAYE 201

DB 230 RADYNGNNITIGENL-----RKTGDLKT--IAELTEEARKQELV----- 268

QY 202 QNPQHIEDLEK--VAVEQLTGHSSVLEELVOLVDRKNIDISIKY-----DPRKDEY 253

DB 269 QNRLQVLEKPKDKMEIEELCSYKS--ELNQLMEKEKNOCKHYREINAIQERTMSHI 325

QY 254 FANRVITDDIELLKLIAFLPEDITLKGHDNDLONIGIKRYK--EPLD--SSPTQWE 309

DB 326 ---QKIVDHEKRLKLESRRKLEIKCNELAKREYVHNTERKLSLEDIENAKSKSSIE 382

QY 310 LRAFWAVHFSITADRIDDIILKYIVDSMNHGDAKSLREELAEITAEIKIYVIAQAEI 369

DB 383 LAA-----MEQKADAEVVKLADQ-----REELHEKIRLERQROQALEYEV 427

QY 370 N-----KHLSSGGINHDSINLMQNLVYTDSEIFKASAEKILEKMPQTTIQ 420

DB 428 EOLKQGLNVRKMAHSAQSGDAEV--VKEVDIIFKDL-----GKEKQALDLKFNQTLIL 478

QY 421 VDSEKKIYVIRKDELSENRRTGALNKLNSYSYNKDNNEHSHFATTCSDSRPLDYS 480

DB 479 ---RERR---TNDELQAEKE---LVNIMKEWNTNIGVRMGEVLT-----KPEYDAMQ 523

QY 481 QKTQQLSDITSRFNSAIEALNRFIOKID 508

DB 524 QKYYCQ-QDVEDRAVEVILQMEHYLKDSD 550

RESULT 13

C71622

hypothetical protein PFB0145c - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 07-May-1999

C:Accession: C71622

R:Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.

; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H

Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600; MUID:99021743

A:Accession: C71622

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1979 <GAP>

A:Cross-references: GB:AE001375; GB:AE001362; NID:g3845105; PID:g3845107; TIGR:PFB014

C:Genetics:

A:Experimental source: clone 3D7

A:gene: PFB0145c

Query Match

Best Local Similarity 20.8%; Score 137; DB 2; Length 1979;

Matches 95; Conservative 71; Mismatches 135; Indels 156; Gaps 20;


```

OY      120 SODNNHOFTKVYGKSPDSDISPKNGENLVGDVVLATGSGDFPVSIGSGKGKLAA 179
       : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      98 NNDNNNEN-----NDDNNNFNNYSDDSIKNIIHKHNDLENOQLKD--TKLSISLSKIY- 145
OY      180 GKYTDAVTAVTSNOEFMIRAEQONPHFIEDLKRVRVEQLTGHGSVVLEEVLQVKOKNI 239
Db      150 -----NYESKIEBELK-----ELKE-VKOKNI 170
OY      240 DISIKYDPKRDSEVANRVITDDIELLKLAFLFPEDTLKGGHYDNOLNQNKREKFE 299
Db      171 DNNYEKKLEKEDE---VKOKIDMLNE-----KENILOKEJLDIN----KEEKKI 214
OY      300 LESSPNTOMELRAFMAYMHSLTAADRIDDDLKYVDYSNKHGDANSKULREELAEIETABL 359
Db      215 NEKSKNIKKKEEPHNIEKEYLEKKNRERETISIEIDIKKH---LEKLIETIKEKKEDL 270
OY      360 KIYSVIOAELINKHLSS-----SGTIINHDSINMLDKNLXYGIDEELFFKASAEIKYL 411
Db      271 E-----NLNKKLLSKENYALKELKGCYVERKETNETINSLANDN-----IIEKKRYKLL 315
OY      412 E-KMPOTTIOVGSEK-----KITV 429
Db      316 EYELBEKNKOJDLNKBDEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKE 375
OY      430 SIR-----DELGESENKRTGALGNLNKSYSYNKDNNELSHRPATTCSDKSRPLNDL 478
       ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      376 SIKREMDIEKRENFLHMEQO-----LKDLKNSFY--KNNNOLKVY--KCEIKRN-LKTE 425
OY      479 VSOKTLOLSDTSRFNSAIETALNELFOKYDSVWORLL 515
Db      426 LEKKEKEKLDIE--NVSKIEINKLINLOLEKEKEOIL 459

```

RESULT 14

RAD50 protein yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein N0872; protein YN250w
C:Species: *Saccharomyces cerevisiae*
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 10-Dec-1999
C:Accession: S05808; S63223
R:Alani, E.; Subbiah, S.; Kleckner, N.
E:Genetics 122, 47-57, 1989
A:Title: The yeast RAD50 gene encodes a predicted 153-kD protein containing a putine nuc
A:Reference number: S05808; MUID:89276917
A:Accession: S05808
A:Molecule type: DNA
A:Residues: 1-1312 <AL>
A:Cross-references: EMBL:X14814; NID:q4272; PIDD:CA32919.1; PID:q4273
R:Sen-Gupta, M.; Guelendiner, U.; Behnhauer, J.; Fiedler, T.; Hegemann, J.H.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63220
A:Accession: S63223
A:Molecule type: DNA
A:Residues: 1-1312 <SEN>
A:Cross-references: EMBL:Z71526; NID:q1302292; PIDD:CA96157.1; PID:e239641; PID:q1302292
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:RAD50; MIPS:YN250w
A:Cross-references: SGD:S0005194; MIPS:YN250w
A:Map position: 14L
C:Superfamily: RAD50 protein
C:Keywords: Arp; coiled coil; DNA repair; meiosis; nucleus
E:177-471/Region: heptad repeats
E:743-995/Region: heptad repeats
E:40/Binding site: Arp (lys) #status predicted

	Query Match	5.13;	Score 136;	DB.1:	Length 1312;
	Best Local Similarity	21.5%;	Pred. No. 14;		
	Matches	96;	Conservative	79;	Mismatches 178; Indels 94; Gaps
Oy	124	NNHQTTVICKRDSRDPISPKYGNGENVGDVVLATGSDQFFVRISGSGKGKLAAKKT	103		
	:	:	: :	: :	

```

Dh 225 NIHQLOKIDQYNEEVAESESQINEITEKSDK--LFKSNQDF--QKILSKVENLKNWK-- 278
QY 184 DAVTVYVSNQEFMIRAVEQNPQHFIEDLEKVRVBOULTGHGSSVLEELVOLVKRKNIDIST 243
Dh 279 ----LSISDQ--VKRLSNSID--IIDLSPQIDQNLILANSFKVLMKNNQRLDLETDISS 329
QY 244 KYDPKRSSEVFANRVITDIDIELLKILATVLPEDTILKGGHYNOLONGAKRKVEPLESS 303
Dh 330 LKPROSSLQSLNSLILROGEL-----BAGEYEEKNNHLSLSLKEAFQHRFOGLSNI 382
QY 304 PNTQW-----ELRAFMAVMHFSILADRID-----DILK-VYVDSMN--HHG 342
Dh 383 ENSMAQVNHMEQOFKAFISQDILT-DTIDFAPADIQLKTNLSDLKISTVDSQNNLEYNK 441
QY 343 DAREKLEELAEIETAEIKIYSVIOAEINKLSSSGTINIHDKSINIMDKULVGYTDEIF 402
Dh 442 KDRKRLIHDEEELAEIKISF-----KLSYQDLSINHELEMLKTYKKELOSWESENIT 493
QY 403 KASAEYKILEMPOTTIYOVGSEKKIYSIDFISGEKKRF--GALGNLKNSSYVNDN 459
Dh 494 P-----KLNQRIEKKNMEMIILENOQIEKFODRIMKTNQOADLYAKGLIKKSINTKID-- 546
QY 460 ELSHFATTCSDKSR-----PLN-----DL-----VSQKTYQLSDITSRFN 494
Dh 547 ELQVITEKLONSDSIRQVFFPLTOEFOQADLEMOFKLFINQKINIAINNNKMHLEDRYT 606
QY 495 SAIEALNRFIQYDVSVMQRLIDTSGR 521
Dh 607 NALYNLN-----TIKRDLODNNKSK 626

```

RESULT

rhophtry protein - Plasmodium yoelii
 C:Species: Plasmodium yoelii
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T28677
 R:Keen, J., Sinha, K., Brown, K., Holder, A.
 Moll. Biochem. Parasitol. 65, 171-177, 1994
 A:Title: A gene coding for a high molecular mass rhophtry protein of Plasmodium yoelii
 A:Reference number: 220508; MUID:95021522
 A:Accession: T28677
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2269 <KEE>
 A:Cross-references: EMBL:L27838; NID:g457145; P1D:g457146; PIDN:AA21304.1

Query Match	5.1%	Score 136,	DB 2:	Length 2269;
Best Local Similarity	22.1%	Pred. No. 29;		
Matches 89;	Conservative 80;	Mismatches 141;	Indels 92;	Gaps 22;

[illegible]

Db 888 K--NINDELNSEK---YIKKIEGDLSEKRSKIN---STDDKDIDECIKININVLKR 938
QY 481 OKTQJSDITSRPSAIE---ALNRF--IQYDSVMORLLD 516
Db 939 NINDETINITNHFKNABEYNNKIIVLSNFNINTEMADNKSQYILLE 980

Search completed: August 22, 2000, 16:50:01
Job time: 1047 sec

QY 376 SGTINHDKSNIMADKLYGTDEEIFKASAEYKILEKMPOTTIOVDGSEKKIYSIDFL 435
 DB 181 SGTINHDKSNIMADKLYGTDEEIFKASAEYKILEKMPOTTIOVDGSEKKIYSIDFL 240
 QY 436 GSEKRTGALGNLKNKSNYKDNNEISHFATTCSDKSRPLNDLYSOKTTOLOSDITSRFS 495
 DB 241 GSEKRTGALGNLKNKSNYKDNNEISHFATTCSDKSRPLNDLYSOKTTOLOSDITSRFS 300
 QY 496 AIEALNRFIOKYDSVMORLLDDTSGK 521
 DB 301 AIEALNRFIOKYDSVMORLLDDTSGK 326

RESULT 2

LCRV_YERPS STANDARD: PRT: 326 AA.
 ID LCRV_YERPS STANDARD: PRT: 326 AA.
 AC P23994;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN V).
 GN LCRV.
 OS Yersinia pseudotuberculosis.
 OG plasmid pIB1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-XPIII;
 RX MEDLINE: 91154114.
 RA Bergman T., Hakansson S., Forsberg A., Norlander L., Macellaro A., Baeckman A., Boelln I., Woll-Waltz H.;
 RT "Analysis of the v antigen lcrGV-yopBD operon of Yersinia pseudotuberculosis: evidence for a regulatory role of LcrV and LcrV."
 RL J. Bacteriol. 173:1607-1616(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97162308.
 RA Roggenkamp A., Gelfer A.M., Lettritz L., Kessler A., Heesemann J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-recombinant V antigen is dependent on polymorphism of v antigen."
 RL Infect. Immun. 65:446-451(1997).
 CC -1- FUNCTION: INVOLVED IN CA(2+) REGULATION OF YOP EXPRESSION, WHICH
 CC INCLUDES THE EXPORT PROCESS.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M57893; AAA27645.1; -;
 DR EMBL: X96802; CAA65594.1; -;
 DR PIR: B37314; B37314.
 KW Plasmid; Antigen; Virulence.
 SO SEQUENCE 326 AA; 37336 MW; 2FD9455DAFF48C06 CRC64;

Query Match 59.6%; Score 1590; DB 1; Length 326;
 Best Local Similarity 96.6%; Pred. No. 2,7e-75;
 Matches 315; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 196 MIRAYBNPOHFEDLEKXVEOULTGHSYVLELVOLVMDKNIDISIKIDPRKDSVFA 255
 DB 1 MIRAYBNPOHFEDLEKXVEOULTGHSYVLELVOLVMDKNIDISIKIDPRKDSVFA 60
 QY 256 NRYITDIELKKTILAFLEPDTILKGHYDNOLNGIKRYKELESPTQWELRAFMA 315

DB 61 NRYITDIELKKTILAFLEPDTILKGHYDNOLNGIKRYKELESPTQWELRAFMA 120
 QY 316 VMPSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKIYSVIOAEINKHLS 375
 DB 121 VHSFLADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKIYSVIOAEINKHLS 180
 QY 376 SGTINHDKSNIMADKLYGTDEEIFKASAEYKILEKMPOTTIOVDGSEKKIYSIDFL 435
 DB 181 SGTINHDKSNIMADKLYGTDEEIFKASAEYKILEKMPOTTIOVDGSEKKIYSIDFL 240
 QY 436 GSEKRTGALGNLKNKSNYKDNNEISHFATTCSDKSRPLNDLYSOKTTOLOSDITSRFS 495
 DB 241 GSEKRTGALGNLKNKSNYKDNNEISHFATTCSDKSRPLNDLYSOKTTOLOSDITSRFS 300
 QY 496 AIEALNRFIOKYDSVMORLLDDTSGK 521
 DB 301 AIEALNRFIOKYDSVMORLLDDTSGK 326

RESULT 3

CAFI_YERPE STANDARD: PRT: 170 AA.
 ID CAFI_YERPE STANDARD: PRT: 170 AA.
 AC P26948;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE F1 CAPSULE ANTIGEN PRECURSOR.
 GN CAFI.
 OS Yersinia pestis.
 OG plasmid pPfr.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 91099503.
 RA Galyov E.E., Smirnov O.Y., Karlishov A.V., Volkovoy K.I.,
 RA Denesjuk A.I., Nazimov I.V., Rudtsov K.S., Adiravov V.M.,
 RA Dalvadnyaz S.M., Zay Yalov V.P.;
 RT "Nucleotide sequence of the Yersinia pestis gene encoding F1 antigen and the primary structure of the protein. Putative T and B cell epitopes."
 RL FEBS Lett. 277:230-237(1990).
 CC -1- SUBCELLULAR LOCATION: CAPSULE.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X61996; CAA43966.1; -;
 DR PIR: S13008; S13008.
 KW Plasmid; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 170
 FT DOMAIN 100 150
 FT
 SO SEQUENCE 170 AA; 17666 MW; 9AC87796A0BA67D9 CRC64;
 POTENTIAL.
 F1 CAPSULE ANTIGEN.
 CONTAINS POTENTIAL ANTIGENIC DETERMINANTS
 THAT MAY STIMULATE T-CELLS.

Query Match 31.9%; Score 850; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1,1e-37;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISSYIALALGATATANADLSTATATLVPARITLYKKGAPITMDNGNIDT 83
 DB 1 MKRISSYIALALGATATANADLSTATATLVPARITLYKKGAPITMDNGNIDT 60
 QY 84 ELLVGLTLGGYGTGTTSTSVNFTDAAGDMYLTFTSOGGNNQFTTKYIGKRSRFDIS 143
 DB 61 ELLVGLTLGGYGTGTTSTSVNFTDAAGDMYLTFTSOGGNNQFTTKYIGKRSRFDIS 120

```

OY 144 PRVNGENLVGDVVLATGSGDFVFRSISGSGKGLAAGRYTDAVTYSNO 193
DB 121 PRVNGENLVGDVVLATGSGDFVFRSISGSGKGLAAGRYTDAVTYSNO 170

RESULT 4
RBP2_PLAVB STANDARD; PRT; 1251 AA.
AC 000799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RETICULOCTE BINDING PROTEIN 2 (FRAGMENT).
GN RBP2.
OS Plasmodium vivax (strain Belen).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92315338.
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
RT merozoites."
RL Cell 69:1213-1226(1992).
CC -1- FUNCTION: INVOLVED IN RETICULOCTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCTE CELLS.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch.)
CC
DR EMBL; M88098; AAA29744.1; -
KW Malaria; Receptor; Membrane.
FT NON TER 1 1
FT NON TER 1 1
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C740A4C572 CRC64;

Query Match
Best Local Similarity 6.2%; Score 164.5; DB 1; Length 1251;
Matches 131; Conservative 104; Mismatches 22; Indels 189; Gaps 33;

OY 8 HHHHSHGHIDDD-----KHKKISVYIALFETIATANAADL----- 47
DB 402 HNYDNAEDELNDKSTNAKVLETNLESVKH--NLSEITINIKGGEKITYSKAKDIMOKIX 459
OY 48 -TASTATATL-----VEPARL-----ILTYEGAPITIM 76
DB 460 AISENTAKTEKVKVDOSNYVNLQITERNLIVTEKNRLNGIDSTITNIEGA--LKE 517
OY 77 DNGNIDELVLGTLG---GYKGTSTSVFTDAAGDPMVLEFTSDGNNHQTTKVI 133
DB 518 SKGNVEIFLEKLEIGKRRKIKADITKKSINSTVGNFSLENNF----- 562
OY 134 GKSDHDPDISPVNG-ENLVGDVVLATGSGDFVFRSISGSGKGLAAGRYTDAVTYSN 192
DB 563 --DLNQYFNKNINDEYKMKGEIYNEPFGSLNKISENL--RNASENTSYNNAKTLRLLEA 618
OY 193 QEFMIRAY--EONPOHETEDLEKV-----RYEOLT--GHGS--S 225
DB 619 QKEKYNLANKKEEANKYLYDVAKVESFRIFENKRESLDKINEMIKKEQLTVNEGHGNYAQ 678
OY 226 VLEELVOLYKNDISIKYPRKDSVPANVITDDIELLKILAYLPE--DTILKNG 283
DB 679 LVENKEIVDENLSDILKQATGKNEITL-----OKITHSTLTKNAKTL--G 723
OY 284 HYDNLQON-GIKRVEF-----LESSPTQWEIR--AFMAVMHFSLTADRID-- 327

```

```

DB 724 HVDTSATYVGKIKTPELALTELLGDAKLTAKTOELKEFSKNNVYLETENMKNTNELDVHK 783
OY 328 --DDLKVIYDSNNHHC--DARSKRLRELAELFAELKIYVIOALNKHSSGTTINIH 383
DB 784 NIDQAYVALEILLAHSEIDTKQDSSKLIEMGNQILKVL--INQYKNKISSISKE 840
OY 384 KSINLMDKNLYGYTDEITFASAEYK-----ILEKPPOTTIO-----VDGSEK 426
DB 841 EAVSVKIGNV-SKHSSELKSLTCSDKSYDNILALEK--QTELQWLRNSFTQEKNTNDS 897
OY 427 KIVSIRK-DELGSENRKRGALGNLK--NSYSYNNDNNELSHFATCSDKSRPLNDVLSOK 482
DB 898 KLEKIKTDFESLKN---ALKTEGEVNALKASDNEH-----HVOSKEPVNPAUSEI 947
OY 483 TTQSLDITSRNSAIEAL-----NRFLOKDSVQWRLDDPT 518
DB 948 EKEETDIDS-LNTALDELKKGRCVEYSRKLIKMDYTKELSDT 991

RESULT 5
REST_CHICK STANDARD; PRT; 1433 AA.
AC 042184; 042228; 057563; 057564;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (CLIP-170).
GN RSV.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98137792.
RA Griparic L., Volosky J.M., Keller T.C. III;
RT "Cloning and expression of chicken CLIP-170 and restin isoforms."
RT Gene 206:195-208(1998).
RN [2]
RP SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
RC TISSUE=PECTORALIS MUSCLE;
RA Griparic L., Keller T.C. III;
RT "Identification and expression of two novel CLIP-170/Restin isoforms
RT expressed predominantly in muscle."
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC CYTOSKELETON (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST FOUR ISOFORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 1 CAP-GLY DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch.)
CC
DR EMBL; AF014012; AAC60344.1; -
DR EMBL; AF020764; AAC60345.1; -
DR EMBL; AF045650; AAC03547.1; -
DR EMBL; AF045651; AAC03548.1; -
DR PFM; PF01302; CAP_GLY. 2.
DR PROSITE; PS00845; CAP_GLY. 1; 2.
KW Cytoskeleton; Microtubules; coiled coil; Alternative splicing.
FT DOMAIN 79 121 CAP-GLY.
FT DOMAIN 144 207 SER-RICH.
FT DOMAIN 235 277 CAP-GLY.
FT DOMAIN 305 332 SER-RICH.
FT DOMAIN 351 1353 COILED COIL (POTENTIAL).

```

```

FT DOMAIN 1414 1427 CCHX-BOX.
FT VARSPLIC 458 492 MISSING (IN SHORT ISOBOM).
FT VARSPLIC 458 492 TOTKEHANIKELBOSLEFKTRADLOLEDETR ->
FT VARSPLIC 803 803 RROKISEDEEN (IN ISOFORM CLIP-170(11)).
FT VARSPLIC 803 803 S -> GGSSKVS (IN ISOFORM CLIP-170(11)).
FT VARSPLIC 458 458 T -> RROKISEDEEN (IN ISOFORM CLIP-170(11+35)).
FT CONFLICT 309 309 K -> R (IN AAC03547).
FT CONFLICT 440 440 E -> V (IN AAC03548).
FT CONFLICT 440 440 E -> V (IN AAC03548).
SQ SEQUENCE 1433 AA: 161026 MW: 5631CE8683498E23 CRC64:

Query Match
Best Local Similarity 20.3%; Score 149; DB 1; Length 1433;
Matches 115; Conservative 100; Mismatches 212; Indels 140; Gaps 28;

QY 38 TATANAADULTATATATATLEPARI-----LTATKE-----GAP 72
DB 140 TAAASATPTSTSTASASASAPALLSPGIPKTSPLAKENSTPSQPSNLSTKASGV 199
QY 73 ITIMDNGNI---DELLVGLTLGKTKGTSTSVNF---TD-AAGDPMTLFTTSODGNN 125
DB 200 SMISEAGSLKKEKREKIGRVLVG---GTRKAGVRFLEGTDFKGEWGEDEDEPLGKN 256
QY 126 HOFFTVIKKSDRDPISP-----KVGEMLVGDVVL-----158
DB 257 ---DGNVAG--TTFVOCORRYGLFAVHVHYTKIGFPSTTPAKKTTVRKVATPALKNS 311
QY 159 ---ANGSDFVRSIGSKGK---AAGKYDAVTVVSNQEFMIRAYEQNPQHFIE 209
DB 312 PSASLSLSSVSVSSKPSRFGTLTETSSRYARKISGTALQ---ALKKQOHIEQ 367
QY 210 ---DLEKVRPEQLGHSSVLELVOLVKNN---IDISIKYPRDSEVFNKRVIT 260
DB 368 LAERLEAEVAKATSHGEIEQELA-LVRDCHDRHAYEMEMKQOLRAVWEAAR-- 423
QY 261 DDEILKRT-----LAFLEPDTILKGHYDNOLOGIK--HYKELESPTQME 309
DB 424 EKEVELLNOLEEKREKVEDLOFREESITRG---DETOTKEHAKIKE-LEOS----- 473
QY 310 LBAFMAVMEFSLADRIDDIKIVYDSNMHNGDASKLREELAEITAKIYSVIOAEI 369
DB 474 -----LFEKTKRADKLQLELEDTRVATVSE-----KSRIMELEHDLALVKEVAEL 519
QY 370 NKHLSSSGINIHDKSINIMDKVLTYTDEETPKASAEKILEKMPOTTIOWDGSKKIV 429
DB 520 KRLSESKHIDVDVTLSDLOE---ISSLOEKMAAKKEHO--REMSSLKKEKESSEALR 575
QY 430 SIKDFLSEKNRKTGALG-NLKNSYSY-NKDNNELSHFATTCSKRSPLMDVYQKTTOLS 487
DB 576 KEKTLTASENENKGENESLKTLDLANKENSDVIELM-----KSK-LESALASHOQAME 629
QY 488 DITSRENSAIEALNRFIOKYDSVMQRL 514
DB 630 ELKVSFNKGVCAQTAEFALKTQMEKV 656

RESULT 6
MSH2_YEAST STANDARD; PRT; 964 AA.
AC P25847; Q12423;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA MISMATCH REPAIR PROTEIN MSH2.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RX MEDLINE; 93093456.
RA Reenan R.A.G., Kolodner R.D.;

```

```

RT RT Isolation and characterization of two Saccharomyces cerevisiae genes
RT encoding homologs of the bacterial Hexa and Muts mismatch repair
RT proteins."
RT Genetics 132:963-973(1992).
RT [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE; 96021609;
RA Zumbstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT "A 29.425 kb segment on the left arm of yeast chromosome XV contains
RT more than twice as many unknown as known open reading frames."
RL Yeast 11:975-986(1995).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE; 93093457;
RA Reenan R.A.G., Kolodner R.D.;
RT "Characterization of insertion mutations in the Saccharomyces
RT cerevisiae MSH1 and MSH2 genes: evidence for separate mitochondrial
RT and nuclear functions."
RL Genetics 132:975-985(1992).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE; 96413311.
RA Alan E.;
RT "The Saccharomyces cerevisiae Msh2 and Msh6 proteins form a complex
RT that specifically binds to duplex oligonucleotides containing
RT mismatched DNA base pairs."
RL Mol. Cell. Biol. 16:5604-5615(1996).
CC -1- FUNCTION: INVOLVED IN POST-REPLICATIVE DNA-MISMATCH REPAIR. BINDS
CC TO MISMATCH-CONTAINING DNA.
CC -1- SUBUNIT: HETERODIMER OF MSH2 AND MSH6.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84170; AAA34802.1; -
DR EMBL; X83121; CAAS8189.1; -
DR EMBL; 274832; CAA99102.1; -
DR PIR; S27433; S27433.
DR SGD; L0001190; MSH2.
DR PFAM; PF00488; MUTS.C; 1.
DR PFAM; PF01624; MUTS.N; 1.
DR PROSITE; P800486; DNA_MISMATCH_REPAIR_2; 1.
KW DNA repair; ATP-binding; DNA-binding.
FT NP_BIND 688 695 ATP (POTENTIAL).
FT CONFLICT 957 964 KYRALL -> EIKRSCCYN (IN REF. 1).
SQ SEQUENCE 964 AA: 108884 MW: 43FED8A640138A84 CRC64;

Query Match
Best Local Similarity 5.4%; Score 145; DB 1; Length 964;
Matches 116; Conservative 98; Mismatches 212; Indels 168; Gaps 27;

QY 47 LASTATATATLEPARITLTYEGAPITIMNG-----NIDEL 85
DB 84 VIVSLQVATLTKLCLDLDDGYK---VEIYDGMWLISASPGNIEQVNEIMNMNIDISI 139
QY 86 LV-----GTLTLAGYRTGTSTSVNFTDAAGDPMTLFTS-----120
DB 140 IIASLKYQWNSODGCIIVAFIDITAKYVGLDLYDNEVSNLESEFLIQGVKCELYOD 199
QY 121 ---QDGNHOTTTRVIG-----KSDRFDISRVNGENLVGDVVLATGSQDF 166
DB 200 LTSNSNSNAEMQVINVIDRCGVVTLTKNSFESEKQVDELDTKLGDIALSLT-PQKYS 258
QY 167 VASIGSKGKLA-----AGKYTDAVTVVSNQEFM-----IRAYEQNPQHF 207

```

```

Db 259 KLISGACNALLGYQLSEDOVGKYE--LVEHKLKEFKMIDASAIKALNLFPGQPNP 315
QY 208 IEDLEKVEVEOLJTGHS--VLEELVOLYKDKNIDISIKY-----DPRKSEVFANR--- 257
Db 316 FGS--NNLAVSGFTSAGNCGKXTSLFQLNHCKNTAGVLLNWMKQPLTNIDIKRDL 374
QY 258 --VITDIELLKILAYLP-----EDTLKGHYDNOLOKIKRVEFLESSPMTQ 307
Db 375 VDYLDOJELKQMLTSEYLPMPIDIRLTKLNKRGNEIDLVL----KYGSKRIPEI- 428
QY 308 WELRAFMAVHFSLTADIDDLKVIYDSNMHGDASKRLRELAELTAEKITS----- 363
Db 429 --VOVFTELEDDSPFEPEVNELVRSWMLAPLSHAEPLSKF--EEVETTVDAEYENNE 485
QY 364 -VIOAEIKR-----HLSSGTINIH-DKSIMLMDKNLYG----- 396
Db 486 FMKVEFNEELCKITSKDLTLDREHSHIDSADLPDPPCKLLENHHLHGWMRLTR 545
QY 397 TDEEIFKASAEKILEKMPQTTIOVGSEKIVSIRKDELGSE--NRTGALG---NLKNS 451
Db 546 NDAKELRKHKKYIELSTVAKAGIFFTKQLKSANETNLOKRYDQGSALVREIINTLT 605
QY 452 YS--YNNKNNELSH-----FATGSDKSRP-----LNDLYSQKTTQSDITSR 492
Db 606 YTPVEFKLSVLALHDVLASFAHTSSYAPIPYIRPKLHPMDESRTHL--ISSR 657

RESULT 7
YD86_SCHPO STANDARD; PRT; 1957 AA.
ID YD86_SCHPO Q10411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYDROTHERMAL 222.8 KDA PROTEIN C1P3.06C IN CHROMOSOME I.
GN SPAC113.06C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
NC Schizosaccharomycetales; Schizosaccharomycetes.
RC [1]
RP SEQUENCE FROM N.A.
RA STRAIN-972;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z70690; CA94624.1; -
RW Hypothetical protein.
SO SEQUENCE 1957 AA; 222785 MW; 3F480CA061719DA CRC64;

Query Match 5.2%; Score 139.5; DB 1; Length 1957;
Best Local Similarity 21.8%; Pred. No. 8.4;
Matches 103; Conservative 74; Mismatches 161; Indels 135; Gaps 21;

QY 145 KVNENLY-----GADV-----VLATGSODFFVRISIGSGKGLAGKTTDAVT 187
Db 1021 KNNGNINISLQTEIKRAENDLQSKLSVSEVENLL--ISSQTNKSLBDR----- 1072
QY 188 VYVSGEEMIRAYEQNPQHFIE--DLEKRVFQLTGHSGSVLEELYQVUKDKNIDISIKY 245
Db 1073 ---TNG---LKYEKNVQKLDEKQKRVLEELTSTKYGKLEENAO--IKDEL--LALRK 1123
QY 246 DPRKSEVFANRYTDDIELLKILAYLPED-----TLKGHYDNO-----LQNGI 293
Db 1124 KSKRQHDICANFV--DDEKESDALEQLTNEKNELIVLSLEQSNMNEALVEERSLANRL 1181

```

```

QY 294 KRYVEFLESSPNTQWELRAFMAVHFSLTADRIDDLIKV-----IYDSNMHNG 342
Db 1182 SDMKRSLSDSNVYSVIRSDLYRVNDELDTLAKRKDLSLTQYSEVQDRODLDSLNGCE 1241
QY 343 DASKRLRELELT-----AEIKIYVIOAEIKNHLSSGT 378
Db 1242 ESFNKYAVSRLNETCKSEIDVPVSEIIDDNFVFNAGNSELSRLTVLSLE--NYLDAFNQ 1299
QY 379 INIHDKSINIMDKNLYGTDEEIFKASAEKILEKMPQTT--TIYDGSKKIV-SIKD 433
Db 1300 VNFKKMEID---NRLLTTTDEFTKVAD--LEKLQHEHDDWLIQGGDEKALKDSEKN 1352
QY 434 FLGSENNKRTGALGNLK-----NSYSYKDNNELSHFATGSDKSRP 474
Db 1353 FLRKEAENTENHSLBEKEETKKEIAELSRLEDNOLATKLNQNDH----- 1401
QY 475 LNDLYSQKTTQSDITSRFSNFAIEALNRFIOKYDSV-----MORLLDDTSRK 521
Db 1402 INOEIRLEKEDVLEKESLILISESLSNQOKESSLLDAKNLELHMLDDTSRK 1454

RESULT 8
RASO_YEAST STANDARD; PRT; 1312 AA.
ID RASO_YEAST P12753;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNA REPAIR PROTEIN RASO50 (153 KDA PROTEIN).
GN RASO50 OR YNL250W OR N0872.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
NC Saccharomycetales; Saccharomycetes.
RC [1]
RP SEQUENCE FROM N.A.
RA STRAIN-RE821;
RX MEDLINE; 89276917.
RA Alani E., Subbiah S., Kleckner N.;
RT "The yeast RASO50 gene encodes a predicted 153-kD protein containing a
RT purine nucleotide-binding domain and two large heptad-repeat
RT regions."
RL Genetics 122:47-57(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE; 97377992.
RA Sen-Gupta M., Gueldeener U., Beinhauer J.D., Fiedler T.A.;
RA Hermann J.H.;
RT "Sequence analysis of the 33 kb long region between ORC5 and SUI1
RT from the left arm of chromosome XIV from Saccharomyces cerevisiae.";
RL Yeast 13:849-860(1997).
RN [1]
RP FUNCTION: IS REQUIRED DURING MEIOSIS IN THE PROPHASE FOR
CC CHROMOSOME SYNAPSIS AND HOMOLOGOUS RECOMBINATION. DURING
CC VEGETATIVE GROWTH IT IS NECESSARY FOR DNA REPAIR. THE PROTEIN MAY
CC BE INVOLVED IN SEARCH FOR HOMOLGY BETWEEN INTERACTING DNA
CC MOLECULES OR CHROMOSOMES. COULD BE A 5'-3' EXONUCLEASE.
CC WORK IN COMPLEX WITH MRE11.
CC -1- SIMILARITY: TO OTHER HEPTAD REPEAT CONTAINING PROTEINS, SUCH AS
CC MYOSINS AND TROPOMYOSINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14814; CA32919.1; -
DR EMBL; X96722; CA65494.1; -
DR EMBL; Z71526; CA96157.1; -
PIR; S05808; BMYIDL.

```

DR SGD: I0001570; RAD50.
 KW Repeat: Heptad repeat pattern; DNA repair; ATP-binding; Coiled coil;
 FT NP_BIND 34 41 ATP (POTENTIAL).
 FT DOMAIN 177 421 COILED COIL (35 HEPTADS).
 FT DOMAIN 743 995 COILED COIL (37 HEPTADS).
 SQ SEQUENCE 1312 AA; 152568 MW; 58A0AA173AC5677E CRC64;

Query Match 5.1%; Score 136; DB 1; Length 1312;
 Best Local Similarity 21.5%; Pred. No. 7.5;
 Matches 96; Conservative 79; Mismatches 178; Indels 94; Gaps 19;

QY 124 NNQFTTKYIGKSRDPEDSPKNGENVDDVYATGSDDFVRSIGSGKLAGKTY 183
 DB NHHQLOTKIDQNEEVESELESQNETTEKSDK--LFRSNDP--OKILSKVENIKYTK-- 278
 QY 184 DAVTVSNQEFMIRAYEONPQHIEDLEKRYEOLTGHSVLEELVOLVKKNIDIS1 243
 DB ---LISDQ---VKRLSNSID--IIDLSKPDQNLANSKYLMKNNQRLDLETDISS 329
 QY 279 ---LISDQ---VKRLSNSID--IIDLSKPDQNLANSKYLMKNNQRLDLETDISS 329
 QY 244 KYDPNRDSEFANRVITDIELLKILAYFLPEDTILKGHYDNOLONGIKRYKEFLSS 303
 DB 330 LKROQSLQSLNSILRQGEI-----EAGREYTKNNHLSLKEAFQHKFGLSNI 382
 QY 304 PNTQW-----ELRAFAVAWHFSLTADRID-----DILK-VYVDSMN--HIG 342
 DB 383 ENSDAQVHNSQKFAFISQDIT-PTIDQFANDIQLEKTNLSDLIKSTIVDSQNLN 441
 QY 343 DASKRLRELAETLAKIYVIOAEINKHLSGGTINIHDKSINIMDKNLVYGTDEIF 402
 DB 442 KDRSKIHSEELAEKLSF-----KSLSTQDSINHELEMLKRYKEKQSWESENI 493
 QY 403 KASAEKIKLEKQPTIQVDGSEKIVSIRKDLGSENKRT--GALGNLKNSSYNKDN 459
 DB 494 P-----KLNQKIEKNEMMILLENQIEKFDRIKMTNQADLYAKILKKSINTKLD-- 546
 QY 460 ELSHFATTSKDSK-----PLN-----VQKTKQLSDITSREN 494
 DB 547 ELQKITEKQNDRIKQVFPDLEQFPRADLEMDFOKLFIMQKNIAINNKMHEDLRRT 606
 QY 495 SAIEALNRFIOKYDYVMORLLDTSGR 521
 DB 607 NALYNL-----TIKDLQDMQSK 626

RESULT 9
 ID MYSD.CAEEL STANDARD: PRT: 1938 AA.
 AC P02567: 019674:
 DT 21-JUL-1986 (rel. 01, Created)
 DT 01-OCT-1989 (rel. 12, Last sequence update)
 DT 15-FEB-2000 (rel. 39, Last annotation update)
 DE MYOSIN HEAVY CHAIN D (MHC D).
 GN MYO-1 OR R06C7.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE: 89178677.
 RA Dobb N.J., Matuyama I.N., Krause M., Karn J.;
 RT "Sequence analysis of the complete Caenorhabditis elegans myosin
 heavy chain gene family.";
 RL J. Mol. Biol. 205:603-613(1989).
 RP [2]
 RP SEQUENCE OF 34-1795 FROM N.A.
 RX MEDLINE: 83273600.
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the Caenorhabditis elegans unc-54
 myosin heavy chain gene are not separated by introns.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).

RN [3]
 RP SEQUENCE OF 115-365 AND 1492-1763 FROM N.A.
 RX MEDLINE: 85201409.
 RA Karn J., Dobb N.J., Miller D.M.;
 RT "Cloning nematode myosin genes.";
 RL Cell Muscle Motil. 6:185-237(1985).
 RP [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Gardner A., McMurray A.;
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
 ELEGANS.
 CC -1- SIMILARITY: THE GLOBULAR HEAD AA SEQUENCE SHOWS A HIGH DEGREE OF
 SIMILARITY WITH THE GLOBULAR HEAD SEQUENCES OF MUSCLE & NONMUSCLE
 HEAVY CHAINS. BY CONTRAST THE ROD SEQUENCE IS LESS CONSERVED, BUT
 THE PEPTIDICITIES OF HYDROPHOBIC & CHARGED RESIDUES, WHICH DICTATE
 THE ALPHA-HELICAL COILED-COIL STRUCTURE, ARE CONSERVED.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: X08065; CA930854.1; -
 DR EMBL: M37232; AAA28119.1; -
 DR EMBL: M37234; AAA28120.1; -
 DR EMBL: Z71265; CA95848.1; -
 DR EMBL: Z71261; CA95848.1; JOINED.
 DR EMBL: Z71261; CA95806.1; -
 DR EMBL: Z71266; CA95806.1; JOINED.
 DR PIR: S02772; MWKML.
 DR HSSP: P08799; LMND.
 DR WORMPEP: R06C7.10; CE06253.
 DR PFAM: PF01576; Myosin_tail; 1.
 DR PFAM: PF00063; Myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Heptad repeat pattern;
 KW Multigene family.
 FT DOMAIN 1 845 GLOBULAR HEAD (S1).
 FT DOMAIN 846 1938 ROD-LIKE TAIL (S2 AND LM DOMAINS).
 FT DOMAIN 846 1170 ALPHA-HELICAL TAIL-PIECE (SHORT S2).
 FT DOMAIN 1171 1938 LIGHT MEROMYOSIN (LM).
 FT NP_BIND 846 1938 COILED COIL (POTENTIAL).
 FT NP_BIND 177 184 ATP (BY SIMILARITY).
 FT DOMAIN 660 682 ACTIN-BINDING.
 FT DOMAIN 764 778 ACTIN-BINDING.
 FT MOD_RES 128 128 METHYLATION (TRI-).
 FT MOD_RES 700 700 ALKYLATION (SH-1).
 FT MOD_RES 710 710 ALKYLATION (SH-2).
 FT CONFLICT 94 94 F -> E (IN REF. 2).
 FT CONFLICT 98 98 A -> R (IN REF. 2).
 FT CONFLICT 377 377 V -> D (IN REF. 4).
 FT CONFLICT 369 390 DV -> GD (IN REF. 2).

FT CONFLICT 391 391 V -> D (IN REF. 4).
 FT CONFLICT 408 408 W -> N (IN REF. 2).
 FT CONFLICT 474 474 Q -> G (IN REF. 2).
 FT CONFLICT 577 577 L -> F (IN REF. 4).
 FT CONFLICT 681 681 I -> N (IN REF. 4).
 FT CONFLICT 1373 1373 S -> D (IN REF. 2).
 FT CONFLICT 1659 1659 E -> Q (IN REF. 3).
 SQ SEQUENCE 1938 AA; 223255 MW; 387399C8F63A4CF4 CRC64;

Query Match 5.1%; Score 135; DB 1; Length 1938;
 Best Local Similarity 19.8%; Pred. No. 14;
 Matches 100; Conservative 89; Mismatches 213; Indels 102; Gaps 19;

QY 76 MNGNID-----TELVGTLGKGTSTSVNFTDAADPMY--LFTSODGN-NHQ 127
 DB 1287 LSNDSNDLAROYELEHILATINRAKTAFSSQVLEAKKAEDLHROEFHACKNLEHE 1346
 QY 128 -----FTTKVIGKDSRDPDIS-----PKVNGENVGD-----D 155
 DB 1347 LPOCHELLENGKIDIOROLSRINSEISQWKAAREGEGLVSELEELKROKRNAMD 1406
 QY 156 VYLTATSGDFEYFRISGSKGKLAAGYTDAYVTYSNOEFMIRAYQNFHF--IEDLER 213
 DB 1407 LDEALSAONKVIETKAKGKLA--ETEDARSVDVRLTLVLALEKORAFDKIVDDMK 1464
 QY 214 VREOOLTGHSVLEELVOLVKDNIDISIKYDPKDSFEFANRV---ITDIELKKI 269
 DB 1465 RAYVDI-----QKEIDATTRDSRNTSTEVFKLRSSMDNLSEQIEFLRRE 1508
 QY 270 LAYFLPE-----DTILKGHYDNOQNGIKRVEKEFLESSPNTQWELRAFV---AVMEF 319
 DB 1509 NRIFSGEIDINEQITQGRTYQEVHKSVRLEQEKDELQHALDEAALAESEKVLRL 1568
 QY 320 SILADIDDDILKVIYDSNNHHGDARSKLRELAELTAEIKTISVQAEINKLSSGII 379
 DB 1569 QLEVOQIRSEIERKIOEKEEFENRKNHORALESIOASLETEAKKAEIARAKKLEET- 1627
 QY 380 NHDKININLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVGSKQIYSIMD-FLGSE 438
 DB 1628 DINOELALDHANKANAVDQKNIK-----KLPOQVKELOGOVDEDERREIENYLALE 1682
 QY 439 NRTGALGNLKNSSYNNKNLSEHPATCSDKSR-----PLNDLYSOKITQLOSDITSRF 493
 DB 1683 KRLALAL-----SSEEDLAH-RIEASDKHKQLEIEQAEKLSMTELLGN----- 1726
 QY 494 NSAIELNRFIOKIVSYMQRLDD 517
 DB 1727 NAALSAMKRRVENEVOIARNELE 1750

RESULT 10
 D7_DICDI STANDARD; PRT; 850 AA.
 AC P54682;
 DT 01-OCT-1996 (rel. 34, Created)
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 01-NOV-1997 (rel. 35, Last annotation update)
 DE CAMP-INDUCIBLE PRESPORE PROTEIN D7 PRECURSOR.
 GN D7.
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Dictyostellida; Dictyostelium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX3;
 RX MEDLINE; 95080502.
 RA Agarwal A., Sloger M.S., Oyama M., Blumberg D.D.;
 RT "Analysis of a novel cyclic Amp inducible prespore gene in
 RT Dictyostelium discoideum: evidence for different patterns of CAMP
 RT regulation."
 RL Differentiation 57:151-162(1994).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY IN THE PRESPORE CELLS;
 CC -1- INDUCTION: BY CAMP.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U25143; AAA73514.1; .
 DR DICTYDB: DD02038; .
 KW Sporulation; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 850 CAMP-INDUCIBLE PRESPORE PROTEIN D7.
 FT DOMAIN 470 475 POLY-GLN.
 FT DOMAIN 555 568 POLY-ASN.
 FT DOMAIN 728 738 POLY-GLN.
 SQ SEQUENCE 850 AA; 95343 MW; 13BA634CCE7AA502 CRC64;

Query Match 5.0%; Score 133.5; DB 1; Length 850;
 Best Local Similarity 19.4%; Pred. No. 5.7; Mismatches 202; Indels 147; Gaps 27;
 Matches 107; Conservative 95;

QY 13 SSGHIDDDDKHKKIISVIALFETIATANAAD-----LTASTATATLVEPARITL 65
 DB 156 TGTGDSKDTTNE-----AIFSSKVSSTDRQESIIQGVALTANDSGATITGLGVSS 210
 QY 66 TYKEGAPITIMNGINDELVLGTLTLLGKYTGTTSTSVNFTDAADPMYLTFTSODGN 125
 DB 211 TAKVGQIT--NGRAQGVVITGDNNTGVGAGVATSAVANVGE--FLGSGRTGSS 265
 QY 126 HQFTTKVIGKDSRDPDISPKVNGENVGDDVY-----LATGSDDFVRSIGSGKLA--- 178
 DB 266 SAGT--VG-----NVISDSTYSGIKIASNGNSLSSETIGTIGDTLHTF 307
 QY 179 AKRYTDAYVT-----VSNQEFMIRAVEQNPOHFIEDLEKRYEOLTG 221
 DB 308 AG--TDSVGVTFPHILITKTFNLIAAGKFNSDQYIDKSGTIPS--QDNEIKKRLDGA 362
 QY 222 HSSVLEF---LVOLVKDNIDISIKYDPKDSFEFANRVITDIELKTLAYFLPED 277
 DB 363 H--QOLQESPAIYQSMKSEDL-----KNIDDEVIRNTLAKEMQIORENDQIGQONED 413
 QY 278 -----TILKGHYDNOQNGIKRVEKEFLESSPNTQ--WELAFAYVMEFSLADRI 326
 DB 414 KQQLIDLQNRPEGLYKNQ-----QDLKQ--ERRANQOELINELNLQEOQEYELLDDL 466
 QY 327 DDDILKVIYDSNNHHGDARSKLRELAELTAEIKTISVQAEINKLSSSGTINIHDKSI 386
 DB 467 YDE-----QOOQOPQVSNKQOLQ--QOINSPEDIQYOLN-HLNOFPDDVYHNDOT 514
 QY 387 NLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVGSSEKRIYSIKDFLGSNKRKGALG 446
 DB 515 EELKDDVYNFNDQINNGQFENN--EFPD-----LNDANDNEQVNA 555
 QY 447 NKNSSYNNKNLSEHPATCSDKSRPLNDLYSOKITQLOSDITSRFNSAIALNRFIOK 506
 DB 556 NNNNNNNNNNNN---FRV--DKSK-----KSAQOVEI-----ALENERLYLOE 594
 QY 507 YDSVMORLLDD 517
 DB 595 VEDAPERLYEE 605

RESULT 11
 YAF3_SCHPO STANDARD; PRT; 1039 AA.
 ID YAF3_SCHPO
 AC 009857;
 DT 01-FEB-1996 (rel. 33, Created)
 DT 01-FEB-1996 (rel. 33, Last sequence update)
 DT 01-FEB-1996 (rel. 33, Last annotation update)
 DE HYPOTHETICAL 118.6 KDA PROTEIN C29E6.03C IN CHROMOSOME 1.


```

Db 344 VAIGATVVDASAVLATDYKRISFDNNQMOVTRLASNTTFVTP-DANGKAAFGLE----- 397
Oy 218 OLTGAGSSVLEELVOLVMDKNIDISIKYDPRKDEFAFNRVITDIELLKLILAEFLPED 277
Db 398 -LTFGTAVND-----SFTLKPVSDAIVNMDVLTIDEK-----IINASEED 439
Oy 278 TILKGHYDNDOLONGIKRYKEFLSSPNTQWELRAFMAVHPSLTADRIDDILKVIYDS 337
Db 440 ---AGDSDN--RNGQALLD--LQSNKRYVGAKSP----- 467
Oy 338 NMHGDASKLEELAEITAEIKITYVIOAEI-----NKHLSSTGINTIHDKSILMDKN 392
Db 468 ---NDAASIVSDIGNKTATLTKTSSATQGNVVTQLSNOQOISIG-VNIDEYGNLORPO 522
Oy 393 LYGTDEELFKAS 405
Db 523 OYTLNMOVLQTA 535

RESULT 13
XLS_YEAST STANDARD: PRT: 1803 AA.
ID XLS_YEAST STANDARD: PRT: 1803 AA.
AC P47024; P87192;
DC 01-FEB-1996 (Rel. 33, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DE 15-FEB-2000 (Rel. 39, Last annotation update)
DN TRANSPOSON TY4 207.7 KDA HYPOTHETICAL PROTEIN.
GN TY4B OR YTL113W OR J0780.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
RN Saccharomycetaceae; Saccharomycetes.
RP [1]
SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE: 97103775;
RA Czaplepruch C., Kordec E., Pujol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including URA2 (5' end), TRK1, PDS2,
RT SP10, GCN14, REP1, PHO86, NCA3, ASF1, CCT7, GEF3, and two tRNA genes,
RT three remnant delta elements and a Ty4 transposon.";
RL Yeast 12:1471-1474(1996).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 249389; CA889409.1; -
RN Transposable element; Hypothetical protein.
KM SEQUENCE 1803 AA; 207693 MW; 16DCD7284AD52D3 CRC64;
SQ

```

Query Match 4.9% Score 131; DB 1; Length 1803;
 Best Local Similarity 19.3%; Pred. No. 21;
 Matches 123; Conservative 103; Mismatches 232; Indels 178; Gaps 31;

```

Oy 13 SSGHIDDD--DKRMKISSVIAIALFGTATANADL-----TASTATATL 57
Db 993 NNSHDEKDYDEKVRHPIEIOEKLVGSKNTIKINDENKISIRKKNIGSLINCLASC 1052
Oy 58 VEPARTITLYKGCAPITTMNDGNIDTELLVGLTLAGYTF-----GTSTVNFETDA--- 109
Db 1053 VITDESITNKDES-----MHNAPRELIOEOLKTNHETSPKESIGTINVKFRATNNE 1106
Oy 110 ---AGDPVYLFTSODGNH---OFTTVIGDSRDPISPKVNGENLVGDDVVLATGS 162
Db 1107 ISLKTGSDTSLPKITLESINNHSNDSTNKVEFEKENHPPTEIDIVMSQDTMESNC 1166
Oy 163 ODFFVRSIGSKGKLAGKRYTD-----AVTVIVS---NOE 194

```

```

Db 1167 QD-----GNNIKELKVTDKNVPTDNGTNGVSPRLQNIASGSPVQTVNKSAPLNKE 1217
Oy 195 FMI-----RAYEONPOHFIEDLE-----KYRVEOLTGHGSSVLEELVOLK----- 235
Db 1218 FSSLNKKRRKRHRDKNNSLTSTELERDKKRKRKRVKILPDMETVASPKIRAIYYNEAI 1277
Oy 236 DKNIDISIKYDPR-----KDSEVFANRVITDIELLKLILAEFLPEDTILKGHY 285
Db 1278 SKNPDLKKEHYKQAYHKEIQLNDKMYF-----DVDV--KTSRSRIPNLVPTPTI 1328
Oy 286 DNQLONGIKRYKEFL--SSPNTQWELRAFMAVHPSLTADRIDDILKY---IVDSMN 339
Db 1329 FTKRNGIYKARIYCRGDTQSPDT-----YSVITTESINHNIKIFILMIANNRN 1377
Oy 340 ---HHGDASKLEEL-----AELTAEIKITYVIOAEI-----EINKHLS-SGTI 379
Db 1378 MEMKTLDINHAFLYAKLEELIYIPPHDRRCVVKLNKALYGLKQSPKEMNDHLQOYLNGI 1437
Oy 380 NIDKS-----INLMDKNLY--GYTDEELFKASAEYKI---LEKMPQT-TIQVDSEKKI 428
Db 1438 GLKDSYTPGLYQTEDKMLNAVYDDCVIAASNDQRIDEINLKSNFELKITGTILDD 1497
Oy 429 VSKIDFLGE--NKRIGALGNLKNSTY-----SYNKD-----NNELSHFATTCSDKSRP 474
Db 1498 VLDFTILGMDLYVYNKRLGTIDLTLSFINRMDKRYNEELKIRKRSIPHMSTY---KIDP 1554
Oy 475 LNDVYSKTTQSLDTSRFSNAIEKLN--RTQKYD 508
Db 1555 KRDVLQMSSEEFROGVKLQOLLGELNRYRHRKRYD 1590

RESULT 14
REST_HUMAN STANDARD: PRT: 1427 AA.
ID REST_HUMAN STANDARD: PRT: 1427 AA.
AC P30622;
DC 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DN 15-JUL-1999 (Rel. 38, Last annotation update)
DE RESTIN (CYTOLASTIC LINKER PROTEIN-170 ALPHA-2) (CLIP-170) (REED-
DE STERNBERG INTERMEDIATE FILAMENT ASSOCIATED PROTEIN).
GN RSN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
SEQUENCE FROM N.A.
RC TISSUE-PERIPHERAL BLOOD MONOCYTES;
RX MEDLINE: 92289675;
RA Blibe G., Delabie J., Bruengen J., Richener H., Asselbergs F.A.M.,
RA Carletti N., Sorg C., Odink K., Tarsay L., Wiesendanger W.,
RA de Wolf-Peters C., Shipman R.;
RT "Restin: a novel intermediate filament-associated protein highly
RT expressed in the Reed-Sternberg cells of Hodgkin's disease.";
RL EMBO J. 11:2103-2113(1992).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE: 92405160;
RA Pliere P., Scheel J., Rickard J.E., Kreis T.E.;
RT "CLIP-170 links endocytic vesicles to microtubules.";
RL Cell 70:887-900(1992).
CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC CYTOSKELETON.
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERNBERG CELLS
CC OF HODGKIN'S DISEASE.
CC -1- SIMILARITY: CONTAINS 1 GAP-GLY DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X64838; CAA6050.1; -
DR EMBL: M97501; AAA55693.1; -
DR PIR: S22695; S22695.

DR MIM: 179838; -
DR PFAM: PF01302; CAP_GLY; 2.
DR PROSITE: PS00845; CAP_GLY.1; 2.
KW Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.

FT DOMAIN 78 120 CAP_GLY.
FT DOMAIN 143 204 SER-RICH.
FT DOMAIN 232 274 CAP_GLY.
FT DOMAIN 304 331 SER-RICH.
FT DOMAIN 350 1342 COILED COIL (POTENTIAL).
FT DOMAIN 1408 1421 CCHC-BOX.
FT VARPPLIC 1457 1481 MISSING (IN SHORT ISOFORM).
FT CONFLICT 1059 1069 D->E (IN REF. 2).
SQ SEQUENCE 1427 AA; 160989 MW; 0A4F16DD94254E8 CRC64;

Query Match 4.9%; Score 130.5; DB 1; Length 1427;
Best Local Similarity 20.3%; Pred. No. 16; Mismatches 134; Indels 87; Gaps 17;
Matches 75; Conservative 74;

OY 167 VRSIGSGKGL--AAKYTDVAVTVVSNQEFMIRAYEONPOHFE-----DLEKVEEOL 219
DB 326 VSSRPRTGILFTSSRYARKISGTALQ-----ALKEKQHQHLEQLAEEDLEAEVANA 381
OY 220 TGGSSVLELVQVADKN-----IDISIKYDPKDESEVFNRIYTDIELKRI----- 269
DB 382 TSHVGEIEELA-LARDGHOVLELAKMDQRTWEADR--EKVELLNLEEEKR 437
OY 270 ---LAFLPEDTILKGGHYDNOLONGIK--RYKEFLSSPNTQWELRAFVAVHESLAD 324
DB 438 VEDLOPVEESTTKG---DLETQTKLEHARKIE-LEQS-----LLEFETKAD 481
OY 325 RIDDDILKVIYDSNMHNGDARSKLRELAETLAKIYVIOAEINKHLSSTGTTINHDK 384
DB 482 KLORELEDRVATVSE-----KSRIMLEKEDLALRQVAVELRRLESNNKPADVDVM 533
OY 385 SINDMKNLYGTDEIFAFASAEYKLEKMPOTTIOVDSSEKIVSIKDFLGSEKRTGA 444
DB 534 SLSL-----DEI-----SSLOEKLEVTRTD---HOREITLSKEHFGARE----- 571
OY 445 LGMLKNSYSYKNNDELSPHATGSDKSRPLNDVSOQTTOJSDINRSFAIEALNRFI 504
DB 572 -----THQEKIRALYATATEKLSKENSLSKELHANKENDVIALWMSKLETA---I 620
OY 505 OKYDSVMQRL 514
DB 621 ASHQAMEEL 630

RESULT 15
Y66P_CAEEL
ID YG6P_CAEEL STANDARD; PRT; 530 AA.
AC P90970; P91831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-FEB-2000 (Rel. 35, Last annotation update)
DE HYPOTHEICAL 60.7 KDA PROTEIN T23G11.8 IN CHROMOSOME 1.
GN T23G11.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nemoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Gardner A.;

RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: 281130; CAB03421.1; -
DR EMBL: 284712; CAB03421.1; JOINED.
DR EMBL: 284712; CAB06546.1; -
DR EMBL: 281130; CAB06546.1; JOINED.
DR WORMPEP: T23G11.8; CE14094.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 505 525 POTENTIAL.
SQ SEQUENCE 530 AA; 60716 MW; AB4376D10C891FA3 CRC64;

Query Match 4.9%; Score 130; DB 1; Length 530;
Best Local Similarity 21.4%; Pred. No. 4.7;
Matches 110; Conservative 84; Mismatches 211; Indels 108; Gaps 24;

OY 40 ATANAD-----LRASTATATLVEPARITLYKESGAPITINDGNIDELLVTLT 91
DB 7 AIONMADSAPRRKSSSTASVSATIIIMPEHVITEFGRRAP--SLPSRAQSD---YTGAGM 62
OY 92 LGG--YKGTGTSVAVFTDAAGDPVYLFETSDQGNHQTAVICKDSDFPISRVKNE 149
DB 63 RGRNRTEQSSSTHQPGVTAASEMTVINSKEVAN-----AKPDEDTISQSSVGR 114
OY 150 NLVGDVVLATGSGQDFVRSIGSGKGLAAGYTDVAVTVVSNQEFMIRAYEONPOHFE- 208
DB 115 SVAS-----IRITSE--FAGNDSSPNKRSQLFADKQSIYVLKTKLSEVEKFEKRSQDYE 168
OY 209 -----EDLEK-YRVEDLTGHGSSVLEL-----VOLYKDKRIDISIKYDPKDESEVFN 256
DB 169 MKAEMKLEKRVENKVSSEHSDLSQELKLAROKADQKEKAVE--ECNMHKKRTYGL 225
OY 257 RYITDIELKILAYFEDTILKGGHYDNOLONGIKRVKEFLSSPNTQWELRAFVAV 316
DB 226 EELRAVVEQLR--LAKFNLENKKEPEDEYKNAQK-ILTAKELVESLSKSEGGIGSSDRP 282
OY 317 MHESLTAD-----RIDDDILKVIYDSNMHNGDARSK--LRELAETLAKI-YSVIOAE 368
DB 283 VHL-LQAEVEEIRVERDLKADLES-----AQLVYILRSMEELEAQIRLOLSQ 335
OY 369 INKHLSSSTINHDKSTLMKNL-----YGYDEELTFKASAEY--KILEKMPOTTIQ 420
DB 336 KRTHEEKOT--WDSTIGLNEKEVCSRIENEFKOEKKRGGDOHOSMLEK----- 385
OY 421 VDGESEKIVSINDFLGSEKRTGALGNLKNYSYKNNDELSPHATGSDKSRPLNDVLS 480
DB 386 -----ENELKRTISDLRSKRRDQTOE-----DDGSSQLTDLIL 420
OY 481 OKTTOISDITSR---FNSAIEALNRFIOKDYDV 510
DB 421 OKOQOLEDVLRNNQVLRNVALERLQKATNRETSTI 453

Search completed: August 22, 2000, 16:51:53
Job time: 294 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 16:46:39 ; Search time 46.06 seconds
(without alignments)
784.261 Million cell updates/sec

Title: US-08-699-716a-2

Sequence: 1 MGHNNHHHHSSGHIDDD.....RFIOKDYVMQRLDPTSGK 521

Scoring table: BLOSUM62
Gapop: 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_12:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1646	61.7	326 2	068697 yersinia pe
2	1557	58.4	324 2	087495 yersinia en
3	1552.5	58.2	334 2	056853 yersinia en
4	1549.5	58.1	334 2	056896 yersinia en
5	1537	57.7	325 2	056880 yersinia en
6	1532	57.5	324 2	056882 yersinia en
7	1531	57.4	324 2	056846 yersinia en
8	1527	57.3	324 2	056879 yersinia en
9	1527	57.3	324 2	056879 yersinia en
10	150.5	56.6	1116 3	030527 pseudomonas
11	149	56.6	2166 2	059247 saccharomyc
12	147	55.4	785 13	090763 gallus gall
13	144	55.4	1115 3	094488 schizosacch
14	139.5	55.2	1928 5	097291 plasmodium
15	139	55.2	924 5	015738 dictyostell
16	137	55.1	1979 5	096133 plasmodium
17	136	55.1	603 4	091607 homo sapien
18	136	55.1	2269 5	026223 plasmodium
19	135	55.1	674 2	09XD51 moraxella c

20	135	5.1	1173 5	P92199 caenorhabdi
21	135	5.1	1510 5	Q22699 caenorhabdi
22	134	5.0	1939 5	Q25662 plasmodium
23	133.5	5.0	601 5	Q76379 echinococcu
24	133.5	5.0	2401 5	Q26216 plasmodium
25	133	5.0	795 5	Q91135 dictyostell
26	132	5.0	168 4	Q16256 homo sapien
27	132	4.9	2606 12	Q36414 alcelaphine
28	131	4.9	1803 3	P87192 saccharomyc
29	129	4.8	510 2	Q51457 borrelia bu
30	129	4.8	694 2	Q92802 chlamydia p
31	129	4.8	819 5	Q9XYP7 dirosophila
32	128.5	4.8	1730 10	Q04454 arabidopsis
33	128	4.8	1542 5	Q97299 plasmodium
34	127.5	4.8	586 2	Q45598 bacillus su
35	127.5	4.8	958 5	Q9XTM0 caenorhabdi
36	127	4.8	576 2	Q54407 moraxella c
37	126.5	4.7	1955 5	Q61308 parascaris
38	126	4.7	552 2	Q9XKFE helicobacte
39	125.5	4.7	799 11	Q54800 rattus norv
40	125	4.7	1173 3	Q13313 schizosacch
41	124.5	4.7	761 13	P79793 gallus gall
42	124.5	4.7	1231 5	Q24098 dirosophila
43	124	4.7	523 2	Q31304 borrelia bu
44	124	4.7	858 1	Q47968 haloflexax s
45	124	4.7	1418 3	Q12267 saccharomyc

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	326 AA.
068697	068697	01-AUG-1998 (TREMBLrel. 07, Created)		
DT	01-AUG-1998 (TREMBLrel. 07, last sequence update)			
DR	01-MAY-1999 (TREMBLrel. 10, last annotation update)			
DE	V ANTIGEN.			
GN	LCRV.			
OS	Yersinia pestis.			
OG	Plasmid pCD1.			
OC	Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Yersinia.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-KIM:			
RA	HU P., ELIOTT J., MCCREADY P., SKOMRONSKI E., GARNES J.,			
RA	KOBAYASHI A., CARRANO A.V., BRUBAKER R., GARCIA E.;			
RL	Submitted (MAR-1998) to the EMBL/Genbank/DBD databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-KIM5:			
RX	MEDLINE: 98427122.			
RA	PERRY R.D., STRALEY S.C., FETHERSTON J.D., ROSE D.J., GREGOR J.,			
RT	BLATTNER F.R.;			
RT	"DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of			
RT	Yersinia pestis KIM5."			
RL	Infect. Immun. 66:4611-4623(1998).			
DR	EMBL: AF053946; AAC65274.1; -			
DR	EMBL: AF074612; AAC65799.1; -			
KW	Plasmid.			
SO	SEQUENCE 326 AA; 37240 MW; 443BDEDC CRC32;			
QY	196 MIRAYONPOHFIEDLEKRVVOLGTGHSSVLEELVOLVKNRNDISIKYDPRKSEVFA 255			
DB	1 MIRAYONPOHFIEDLEKRVVOLGTGHSSVLEELVOLVKNRNDISIKYDPRKSEVFA 60			

Query Match 61.7%: Score 1646; DB 2: Length 326;

Best local similarity 99.7%: Pred. No. 8.9e-80;

Matches 325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	256	NRVIDDELLKILAYLPEDDTLKGGHYNOLONGIKRYKELESSPUMOWELRAFMA	315
Db	61	NRVIDDELLKILAYLPEDDTLKGGHYNOLONGIKRYKELESSPUMOWELRAFMA	120
Qy	316	VMHSELTADRIDDDILKAYIVDSNMHHGDAKSLREELAEHLTAELDKITYSVJOAEINKHLSS	375
Db	121	VMHFELTADRIDDDILKAYIVDSNMHHGDAKSLREELAEHLTAELDKITYSVJOAEINKHLSS	180
Qy	376	SGTINHKSINLMDKNLXGYTDEEIFAASXYKILTEKMQTTIQUVSGSEKIVYSIDFL	435
Db	181	SGTINHKSINLMDKNLYGTDEEIFAASXYKILTEKMQTTIQUVSGSEKIVYSIDFL	240
Qy	436	GSENRKTCALGNLKNXSYSYNDDNNELSHFATCGDSKSRPLNDLYSOKTTLSDITSRFS	495
Db	241	GSENRKTCALGNLKNXSYSYNDDNNELSHFATCGDSKSRPLNDLYSOKTTLSDITSRFS	300
Qy	496	ALFALNRFIQKYDYSMORLLDDTISGK	521
Db	301	ALFALNRFIQKYDYSMORLLDDTISGK	326

```

RESULT      2
087495
ID          087495      PRELIMINARY;      PRT;      324 AA.
AC          087495;
DT          01-NOV-1998 (TREMBLrel. 08, Created)
DT          01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT          01-NOV-1999 (TREMBLrel. 12, last annotation update)
DE          LCRV PROTEIN.
GN          LCRV.
OS          Yersinia enterocolitica.
OC          Bacteroidia; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC          Yersinia.
NC          [1]
NR          RP
RP          SEQUENCE FROM N.A.
RC          STRAIN-W22703;
RA          IRIARIE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
RT          "Complete sequence of the Yersinia enterocolitica pYV plasmid.";
RL          submitted (JUL-1998) to the EMBL/GenBank/DBSJ databases.
[2]
RP          SEQUENCE FROM N.A.
RC          STRAIN-W22703;
RA          IRIARIE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
RT          "Petalled genetic map of the pYve227 plasmid of Yersinia
RT          enterocolitica serotype O:9.";
RL          Submitted (OCT-1998) to the EMBL/GenBank/DBSJ databases.
DR          EMBL: AF102390; AD16815.1; -.
KW          plasmid.
SQ          SEQUENCE      324 AA: 37296 MM; 7C9A652F CRC32;

```

Query Match	58.4%	Score 1557	DB 2	Length 324
Best Local Similarity	95.4%	Pred. No. 4.3e-75		
Matches 308; Conservative	7	Mismatches 8	Indels 0	Gaps 0

QY	196	MIRAYEQNPHQFIEDLEKVRVQOJLTGGSSVLEETVQVYVNDKNDISIKTDPKRDSEVFA	255
Db	1	MIRAYEQNPHQFIEDLEKVRVQOJLTGGSSVLEETVQVYVNDKNDISIKTDPKRDSEVFA	60
QY	256	NRVITDDIELLKILAYFLPEDTILKGGHYDNOJONGIKRKYKEFLESSPTOMELRAFMA	315
Db	61	DRVITDDIELLKILAYFLPEDTILKGGHYDNOJONGIKRKYKEFLESSPTOMELRAFMA	120
QY	316	VMHSLFLNDRIDDDILKAYVDSNNHHGDASKLKEELAEELTAELKITSVYQALINKHLS	375
Db	121	VMHSLFLNDRIDDDILKAYVDSNNHHGDASKLKEELAEELTAELKITSVYQALINKHLS	180
QY	376	SGTINIHDSINLMDKNLYGTDEEIFKASAEYKILEKMPQTTIYOVDSSEKTIYSINDFI	435
Db	181	SGTINIHDSINLMDKNLYGTDEEIFKASAEYKILEKMPQTTIYOVDSSEKTIYSINDFI	240
QY	436	GSEKRFQALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDIVSOQTLOJSDITSRNFNS	495

```

Db      241 ESEKRTGALGNLDOSVSYNNDNNELSHFATTCSDKSRPLNDLYSQKTTGLDSDITSRFS 300
      :
Qy      496 AIEALNRFIQKTYDSVMQRLDDT 518
      |||||
Db      301 AIEALNRFIQKTYDSVMQRLDDT 323

```

RESULT	3			
ID	056853	PRELIMINARY;	PRT;	334 AA.
AC	Q56853			
MC	Q56853;			
DT	01-NOV-1996 (TREMBLrel. 01, Created).			
DT	01-NOV-1996 (TREMBLrel. 01, last sequence update)			
DT	01-NOV-1998 (TREMBLrel. 08, last annotation update)			
DE	V ANTIGEN (FRAGMENT).			
OS	Yersinia enterocolitica.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Yersinia.			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-WA-314;			
RX	MEDLINE; 97162308.			
RA	ROGGENKAMP A., GETIGER A.M., LETTRITZ L., KESSLER A., HEESEKAMP J.;			
RT	"Passive immunity to infection with Yersinia spp. mediated by anti-			
RT	recombinant V antigen is dependent on polymorphism of V antigen";			
RL	Infect. Immun. 65:446-451(1997).			
DR	EMBL; X96796; CAA65588.1.			
FT	NON_TER			
SO	SEQUENCE			
	334			
	334			
	334 AA;			
	38241 MW;			
	71229360 CRC32;			

Query Match	Best Local Similarity	58.2%	Score 1552.5	DB 2	Length 334
Matches	Conservative	93.1%	Pred. No. 7.7e-75	Mismatches 6	Indels 9
					Gaps 1
QY	196	MIRAEQNPNQHIEDLEKRYEQLTGHGSSVLEELVQLVKDKNIDISIXYDPKDSVEFA	255		
Db	1	MIRAEQNPNQHIEDLEKRYEQLTGHGSSVLEELVQLVKDKKIDISIXYDPKDSVEFA	60		
QY	256	NRVYTDIEELKLIAYFLPEDTILKGGHYDQNLONGIKRYKEFLLESSPNTOMELRAFA	315		
Db	61	ERVYDDIEELKLIAYFLPEDAILKGGHYDQNLONGIKRYKEFLLESSPNTOMELRAFA	120		
QY	316	VHFEFLTDRIDDILIKYIVDSMNHGDAKSLRELAELTKEIKYISYQIAEINKHSS	375		
Db	121	VHFEFLTDRIDDILIKYIVDSMNHGDAKSLRELAELTKEIKYISYQIAEINKHSS	180		
QY	376	SGTINIHKSINIMDKNTGYDEEIFKASAEYKILKKAPQTICVD-----GSEK	426		
Db	181	SGTINIHKSINIMDKNTGYDEEIFKASAEYKILKKAPQTITINDELHEGVYIAGAAR	240		
QY	427	KIVSYKDPGLGSENNKKTGALGNLKSYSYKKNDELSHFATTCSDKSRPLINDIVSQRTOL	486		
Db	241	QIVSYKKNLESENKKTGALGNLKSYSYKKNDELSHFATTCSDKSRPLINDIVSQRTOL	300		
QY	487	SDITSRFNSAIEALNRFIOKDYDVQORLLDPT	518		
Db	301	SDITSRFNSAIEALNRFIOKDYDVQORLLDPT	332		

RESULT	4	
056896	056896	PRELIMINARY; PRT; 334 AA.
AC	056896;	
DT	01-NOV-1996 (TREMBLrel. 01, Created)	
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)	
DT	01-NOV-1998 (TREMBLrel. 08, Last annotation update)	
DE	V ANTIGEN (FRAGMENT).	
OS	Yersinia enterocolitica.	
CC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae.	
CC	Yersinia.	
IN	[1]	

RP SEQUENCE FROM N.A.
 RC STRAIN-NCIC;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen."
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96797; CAA65589.1; -
 FT NON_TER 334 334
 SQ SEQUENCE 334 AA; 38227 MW; F23PDED3 CRC32;

Query Match 58.1%; Score 1549.5; DB 2; Length 334;
 Best Local Similarity 92.8%; Pred. No. 1.le-74;
 Matches 308; Conservative 9; Mismatches 6; Indels 9; Gaps 1;

QY 196 MIRAEQNPQHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 255
 DB 1 MIRAEQNPQHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 60
 QY 256 NRVTDDIELLKLILAYFLPDTILKGGHYDNOLONGIKRVEFLLESSPNTQWELRAFMA 315
 DB 61 ERVTDDIELLKLILAYFLPDTILKGGHYDNOLONGIKRVEFLLESSPNTQWELRAFMA 120
 QY 316 VMHSLTADRIDDDILKIVDSNMHGDARSKLRELAELTAEIKIYVIOAEINKHLSS 375
 DB 121 VMHSLTADRIDDDILKIVDSNMHGDARSKLRELAELTAEIKIYVIOAEINKHLSS 180
 QY 376 SGTINIHDSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVD-----GSEK 426
 DB 181 SGTINIHDSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDLEHGVYAGAEK 240
 QY 427 KIVSKDPLGSEKRTGALGNKNSYKKNKNNELSHFATTCSDKSRPLNDVSKTQOL 486
 DB 241 QIVSKNFESEKRTGALGNKNSYKKNKNNELSHFATTCSDKSRPLNDVSKTQOL 300
 QY 487 SDITSRFNSAIEALNRFIOKYDSVMQRLDDT 518
 DB 301 SDITSRFNSAIEALNRFIOKYDSVMQRLDDT 332

RESULT 5
 056880 PRELIMINARY: PRT: 325 AA.

AC 056880;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE V ANTIGEN (FRAGMENT).
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-8081;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen."
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96798; CAA65590.1; -
 FT NON_TER 325 325
 SQ SEQUENCE 325 AA; 37260 MW; DEE88446 CRC32;

Query Match 57.7%; Score 1537; DB 2; Length 325;
 Best Local Similarity 94.4%; Pred. No. 4.8e-74;
 Matches 305; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 196 MIRAEQNPQHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 255
 DB 1 MIRAEQNPQHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 60

QY 256 NRVTDDIELLKLILAYFLPDTILKGGHYDNOLONGIKRVEFLLESSPNTQWELRAFMA 315
 DB 61 ERVTDDIELLKLILAYFLPDTILKGGHYDNOLONGIKRVEFLLESSPNTQWELRAFMA 120
 QY 316 VMHSLTADRIDDDILKIVDSNMHGDARSKLRELAELTAEIKIYVIOAEINKHLSS 375
 DB 121 VMHSLTADRIDDDILKIVDSNMHGDARSKLRELAELTAEIKIYVIOAEINKHLSS 180
 QY 376 SGTINIHDSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDGSEKKIYSINDEL 435
 DB 181 SGTINIHDSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDGSEKKIYSINDEL 240
 QY 436 GSEKRTGALGNKNSYKKNKNNELSHFATTCSDKSRPLNDVSKTQOLSDITSRENS 495
 DB 241 ESEKRTGALGNKNSYKKNKNNELSHFATTCSDKSRPLNDVSKTQOLSDITSRENS 300
 QY 496 AIEALNRFIOKYDSVMQRLDDT 518
 DB 301 AIEALNRFIOKYDSVMQRLDDT 323

RESULT 6
 056882 PRELIMINARY: PRT: 324 AA.

AC 056882;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE V ANTIGEN.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y-96-P;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen."
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96801; CAA65593.1; -
 SQ SEQUENCE 324 AA; 37162 MW; 56842D0A CRC32;

Query Match 57.5%; Score 1532; DB 2; Length 324;
 Best Local Similarity 93.8%; Pred. No. 8.8e-74;
 Matches 303; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 196 MIRAEQNPQHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 255
 DB 1 MIRAEQNPQHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 60
 QY 256 NRVTDDIELLKLILAYFLPDTILKGGHYDNOLONGIKRVEFLLESSPNTQWELRAFMA 315
 DB 61 DRVTDDIELLKLILAYFLPDTILKGGHYDNOLONGIKRVEFLLESSPNTQWELRAFMA 120
 QY 316 VMHSLTADRIDDDILKIVDSNMHGDARSKLRELAELTAEIKIYVIOAEINKHLSS 375
 DB 121 VMHSLTADRIDDDILKIVDSNMHGDARSKLRELAELTAEIKIYVIOAEINKHLSS 180
 QY 376 SGTINIHDSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDGSEKKIYSINDEL 435
 DB 181 SGTINIHDSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDGSEKKIYSINDEL 240
 QY 436 GSEKRTGALGNKNSYKKNKNNELSHFATTCSDKSRPLNDVSKTQOLSDITSRENS 495
 DB 241 ESEKRTGALGNKNSYKKNKNNELSHFATTCSDKSRPLNDVSKTQOLSDITSRENS 300
 QY 496 AIEALNRFIOKYDSVMQRLDDT 518
 DB 301 AIEALNRFIOKYDSVMQRLDDT 323

RESULT 7
 056846 PRELIMINARY: PRT: 324 AA.
 AC 056846;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, last annotation update)
 DE V ANTIGEN.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision: Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y-108-P;
 RX MEDLINE: 97162308.
 RA ROGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96799; CAA65591.1;
 SQ SEQUENCE 324 AA; 37194 MW; 5BD357FF CRC32;

Query Match 57.4%; Score 1531; DB 2; Length 324;
 Best Local Similarity 94.1%; Pred. No. 1e-73;
 Matches 304; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 196 MIRAYEQNHQHTIEDLEKRVQQLTGHSSVLEELVQVKNKIDISIKYPRKSEVFA 255
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 MIRAYEQNHQHTIEDLEKRVQQLTGHSSVLEELVQVKNKIDISIKYPRKSEVFA 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 256 NRVITDDIELKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLSSPNTOMELRAFMA 315
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 NRVITDDIELKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLSSPNTOMELRAFMA 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 316 VNHFSLTADRIDDLIKYIVDSMNHNGDARSKLRELAELTAELKIYVIOAEINKHLSS 375
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 VNHFSLTADRIDDLIKYIVDSMNHNGDARSKLRELAELTAELKIYVIOAEINKHLSS 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 376 SGTINIHDKSINLMKKNLYGYTDEIEFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFL 435
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 181 SGTINIHDKSINLMKKNLYGYTDEIEFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFL 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 436 GSENRRTGALGNLKNYSYKNKDNELSHFATCSQKSRPLNDVSOQTOTSDITSRFS 495
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 ESENRRTGALGNLKNYSYKNKDNELSHFATCSQKSRPLNDVSOQTOTSDITSRFS 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 496 AIEALNRFIOKYDSVQORLLDDT 518
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 301 AIEALNRFIOKYDSVQORLLDDT 323

RESULT 8
 056879 PRELIMINARY: PRT: 324 AA.
 AC 056879;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, last annotation update)
 DE V ANTIGEN.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision: Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97162308.
 RA ROGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96800; CAA65592.1;
 SQ SEQUENCE 324 AA; 37195 MW; ELCB300F CRC32;

Query Match 57.3%; Score 1527; DB 2; Length 324;
 Best Local Similarity 93.8%; Pred. No. 1.6e-73;
 Matches 303; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 196 MIRAYEQNHQHTIEDLEKRVQQLTGHSSVLEELVQVKNKIDISIKYPRKSEVFA 255
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 MIRAYEQNHQHTIEDLEKRVQQLTGHSSVLEELVQVKNKIDISIKYPRKSEVFA 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 256 NRVITDDIELKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLSSPNTOMELRAFMA 315
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 NRVITDDIELKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLSSPNTOMELRAFMA 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 316 VNHFSLTADRIDDLIKYIVDSMNHNGDARSKLRELAELTAELKIYVIOAEINKHLSS 375
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 VNHFSLTADRIDDLIKYIVDSMNHNGDARSKLRELAELTAELKIYVIOAEINKHLSS 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 376 SGTINIHDKSINLMKKNLYGYTDEIEFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFL 435
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 181 SGTINIHDKSINLMKKNLYGYTDEIEFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFL 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 436 GSENRRTGALGNLKNYSYKNKDNELSHFATCSQKSRPLNDVSOQTOTSDITSRFS 495
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 ESENRRTGALGNLKNYSYKNKDNELSHFATCSQKSRPLNDVSOQTOTSDITSRFS 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 496 AIEALNRFIOKYDSVQORLLDDT 518
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 301 AIEALNRFIOKYDSVQORLLDDT 323

RESULT 9
 030527 PRELIMINARY: PRT: 294 AA.
 ID 030527;
 AC 030527;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, last annotation update)
 DE PCR.
 GN PCRV.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision: Pseudomonas group;
 OC Pseudomonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-388;
 RX MEDLINE: 98037517.
 RA YAHF T.L., MENDEL-MUELLER L.M., FRIESE M.B., FRANK D.W.;
 RT "Identification of type III secreted products of the Pseudomonas
 aeruginosa exoenzyme S regulon";
 RL J. Bacteriol. 179:7165-7168(1997).
 DR EMBL: AF010149; AAC45935.1;
 SQ SEQUENCE 294 AA; 32283 MW; OFAD2082 CRC32;

Query Match 18.4%; Score 490; DB 2; Length 294;
 Best Local Similarity 39.7%; Pred. No. 5.7e-19;
 Matches 117; Conservative 54; Mismatches 92; Indels 32; Gaps 7;

QY 224 SVLEELVQVKNKIDISIKYPRKSEVFAANRVITDDIELKKILAYFLPEDTILKGG 283
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 26 SAEQELALALSERKIVYAHAGQPLSEAOVL-----KALMLLAANSAPPG 72
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 284 HYDNOLONGIKRVEFLSSPNTOMELRAFMAVNHFSLTADRIDDLIKYIVDSMNH 341
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 73 -----QGLEVLRREVLAQARQPGQMDLREIVAYSLSLNG-RDEDEVIGVYKDVLOTQ 124
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 342 GDARSKLRELAELTAELKIYVIOAEINKHLSSGTTINIHDKSINLMKKNLYGYTDEE 400
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 125 DGRKRALDELKALTAELKIYVIOAEINKHLSSGTTINIHDKSINLMKKNLYGYTDEE 184
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 401 IFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFLGSSENRRTGALGNLKNYSYKNK 460
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 185 RWKDSPEYALSNL-----DTFSGKL-STKDFLSSGPKOSGELKGLSDYEPPEKDNPP 236

```

QY      461 LSHFATCSDKSRLNDLVSQKTQTLSDTSRFNSALREALNRIQKYDSWQRL 515
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
db      237 VGNFATVSDRSRLNDKVNKXTLLNDTSSRYNSAWEALNRIQKYDSVLRDIL 291

```

RESULT	10		
099247			
ID	099247	PRELIMINARY:	PRT: 1116 AA.
AC	099247;		
DT	01-NOV-1996 (TEMBLrel. 01, Created)		
DI	01-NOV-1996 (TEMBLrel. 01, last sequence update)		
DT	01-NOV-1998 (TEMBLrel. 08, last annotation update)		
DE	ORF YOL087C FROM CHROMOSOME XV.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;		
OC	Saccharomycetaceae; Saccharomyces.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN: F11679;		
RX	MEDLINE: 96021609.		
RA	ZUMSTEIN E., PEARSON B.M., KALOGEROPOULOS A., SCHWEIZER M.;		
RT	A 29,442 kb segment on the left arm of yeast chromosome XV contains		
RL	more than twice as many unknown as known open reading frames.";		
RL	Yeast 11:975-986(1995).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	ZUMSTEIN E., PEARSON B.M., KALOGEROPOULOS A., SCHWEIZER M.;		
RL	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: X63121; CAA58192.1; -		
DR	EMBL: Z74829; CAA90099.1; -		
SO	SEQUENCE 1116 AA; 125381 MW; 099CF031 CRC32;		

```
QY      399 EEIFKASAEYKIEKMPQTTIQVDGSEKKIYSIK 4322
      : : : : : : : : : : : : : : : : : :
Db      800 QYMTSSLKY-LTKRLPVTIK-IKASSCPILRVK 8311
```

RESULT	ID	PRELIMINARY:	PRT:	2166 AA.
051465	051465			
AC	051465			
DT	01-JUN-1998	(TrEMBLrel. 06, Created)		
DT	01-JUN-1998	(TrEMBLrel. 06, Last sequence update)		
DT	01-NOV-1999	(TrEMBLrel. 12, Last annotation update)		
DE	HYPOTHETICAL 254.2 KD PROTEIN.			
GN	BB0512.			
OS	<i>Borrelia burgdorferi</i> (Lyme disease spirochete).			
OC	Bacteria; Spirochaetales; Spirochaetaceae; <i>Borrelia</i> .			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE: 98065943.			
RA	FRASER C.M., CASTENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,			
RA	LAHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKET E.K., GWINN M.,			
RA	DOUGHERTY B., TOMB J.-E., FLEISCHMANN R.D., RICHARDSON D.,			
RA	PETERSON J., KERLAUVAGE A.R., QUACKENBUSH J., SALTZBERG S., HANSON M.,			
RA	VAN VOGT R., PALMER N., ADAMS M.D., GOCZYNE J.D., WEIDMAN J.,			
RA	UTTERBACK T., MATTHEY L., MCDONALD L., ARTACH P., BOWMAN C.,			
RA	GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,			
RA	SMITH H.O., VERNER J.C.,			
RT	"genomic sequence of a Lyme disease spirochaete, <i>Borrelia</i>			
RT	<i>burgdorferi</i> ."			
RL	Nature 390:580-586(1997).			
DR	HESBL; AE001153; AAC66876.1; -.			
DR	HSSP; Q57997; IMH.			
DR	TIGR; BB0512.			
KW	Hypothetical protein			
QO	SEQUENCE	2166 AA;	254243 MW;	F28999A0A CRC32;

Db 1235 CYGFINSGEIKAGVENIKN-----HF-DVCIKKVNLLDDIYKVE----- 1277

QY 487 SDITSRNSAIEALNFIQKYSVMORLLDSTG 520

Db 1278 ----NEIHKRIDSLKSIESTFDSIEKMLNDKVS6 1307

RESULT 12

ID 090763 PRELIMINARY; PRT; 785 AA.

AC 090763; 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DE 01-NOV-1999 (Tremblrel. 12, Last annotation update)

OS CHICKEN CADHERIN-7.

OC Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;

OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

RP SEQUENCE FROM N.A.

RC STRAIN-WHITE LEGRORN; TISSUE-BRAIN;

RA MEDLINE; 95309115.

NAKAGAWA S., TAKEICHI M.,

"Neural crest cell-cell adhesion controlled by sequential and

subpopulation-specific expression of novel cadherins."

Development 121:1321-1332(1995).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).

EMBL; D42150; BAA07721.1; .

HSSP; P15116; INCD.

PROSITE; PS00232; CADHERIN; 3.

PFAM; PF00028; cadherin; 5.

PFAM; PF01049; cadherin_C-term; 1.

PRINTS; PR00205; CADHERIN.

Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.

KW SEQUENCE 785 AA; 87171 MW; F27A6881 CRC32;

Query Match 5.5%; Score 147; DB 13; Length 785;

Best Local Similarity 23.0%; Pred. No. 2.6; Mismatches 201; Indels 114; Gaps 26;

Matches 115; Conservative 71; Mismatches 201; Indels 114; Gaps 26;

QY 12 HSSGHIDDDDKHKKI-SSVIAIALFETIATANAADLTAS-----TTATATLVEPARIT 64
 Db 72 HSD -VDKGGSGIKYILISGASISF--ITDENTGDIHAIRLDREQATYTTLRQAADR 127
 QY 65 LTYKESAP-----TTMDNGNIDTELLVGLLGK--TGTTSTSVNFTDAAGPMYL 116
 Db 128 LTRKPYEPSESEFIKIODINDNEPKFLDGYTAGVPEMSPTGTSVQVATD-ADPTY- 185
 QY 117 TFSOGNNHOFITKYIGKSRDFDISPK-----NENLVGDVYVATSGDFEVR 168
 Db 186 -----GNSARVYVSL-QQPYFSEVPKGIKTLPPNDRRAKQOYLVIYAKMD-- 235
 QY 169 SIGSGKGLAGKYTDAVTVVTSNOEFMIRAYQNPQHIEDLEKRVDELTHGSSVLE 228
 Db 236 -VGQNG--LSG--TTSVYTLTD-----VNDNPRFRASYQVNPV-----SLPLA 278
 QY 229 ELVQVYKKNIDISIKYDPKDSVFANRYITDIELKILAYFLPEDTILKGHYDNO 288
 Db 279 SVARAKADADV-----PVAEMEY--KIVDGDGLGVFI-----SVDKD 317
 QY 289 LONGIRKVEFELESSPRTQELAFMAVMA-----FSLTADRIDDILKIVYDSMHHGDA 344
 Db 318 TEGGITTITKEIDFEAKTSTLRIEAMNVDPRFLSLGFSMTYVKIIVEDV----- 371
 QY 345 RSKRLRELAELTALKIYVIOAEINKHSSGTTINHDSSIN-----LMDKNLYGYTD 398
 Db 372 -----DEPVFTS--RLISKVYSEAAKVGTTITVAHDDASNSPVRSIDRN---TD 420
 QY 399 -EEIFRASAERYKLEKMPOTTIOVDSEKKIVSIKDFLGSNKRRTGALNKLKSYSNKD 457
 Db 421 LERYFNIDANSYI-----TTAKSLDRETNVAVNITVYLAESQNPQAIQNGYVAITILLDI 475

QY 458 NNELSHPA----TTCGDKSRP 474

Db 476 NDNAPEFAMEYETTVCEANOP 496

RESULT 13

ID 094488 PRELIMINARY; PRT; 1115 AA.

AC 094488; 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DE 01-MAY-1999 (Tremblrel. 10, Last annotation update)

GN PUTATIVE COILED-COIL PROTEIN.

SPC417.07C.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

RP SEQUENCE FROM N.A.

RC STRAIN-972H-;

RA LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKERT G.;

Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.

EMBL; AF035076; CAA22653.1; .

SEQUENCE 1115 AA; 128468 MW; 6E182EE3 CRC32;

Query Match 5.4%; Score 144; DB 3; Length 1115;

Best Local Similarity 18.6%; Pred. No. 6; Mismatches 180; Indels 258; Gaps 29;

Matches 123; Conservative 100; Mismatches 180; Indels 258; Gaps 29;

QY 16 HIDDDEK-----H-----MKRISVIAIALFETIATANAADLTASTT 52
 Db 182 HDEDDDLQTPKEERNYLSQDAPPTNALSKISDIL-----IPASAKMLDKRN 234
 QY 53 ATATLVEPAR--TTLTKKCAPITMDN----- 78
 Db 235 ALAKEFEESOPGSSLTLRKDA--NVIDNLRKEVFGKLCYLYDQJLNFHQEVDINK 292
 QY 79 GNIDTELLVGLT--TLGKYKGTSTSVNFTDAAGPMYLTFSQDGNHOFITKYGKD 136
 Db 293 QNIDKTLMELOKAVAGE-----KATSGLE 319
 QY 137 SR-----DFDIS--PKVNGENLVGD-----VLTAGSQDF--FVRSIGSGKGLAAGY 182
 Db 320 SRIKPDQSNLSPPSPAPSNLITLQSRYSQALSLETTKRAFAALRKEKSKTNSVGA 379
 QY 183 TDVATVTVNOEFMIRAYQNPQHIEDLEKRVDELTHGSSVLELVQVYKKNIDIS 242
 Db 380 NEDRNV-LSN--MIDNERREKALQELSLRVO-----LSKRYPM----- 417
 QY 243 IKDPKDSVEFANRYITDIELKILAYFLPEDTILKGHYDNOLOGIKRVKEF-- 299
 Db 418 ----PAKNTDERVIELQSNELLRDIS--KONELLARKQENDRL--VQVVELTYA 468
 QY 300 -----LESSPRTQEL--LRAEYAVMHSFLADR-----IDDLIKKIVYDSMN 339
 Db 469 LNSGKNALIVEASSKNELMDSMWMSRMKTQOSIELTFLYQLODIEDYENKMLRMQ 528
 QY 340 HHGDAASKURELAELTAL-----KIYVIOAEINKHSSSG 377
 Db 529 QWRVEDVQLOEYVEITQELDQTEKVLSSKESDDEYEVGKLRTAEAREIEKF--EK 585
 QY 378 TIRIHDSINLMDKNLYGTDEEIFKASAEY----- 408
 Db 586 TIRENESISLREVEKLTID--ETIOLSRVYNDKCHEPELOKRLQTEENNAKEDST 644
 QY 409 ----KILEKMPOTTIOVDSEKKIVSINKDFLGSNKRRTGALNKLKSYSNKD 464
 Db 645 SKTSNLEQLKMTAEVDSLKE-----NEENKQVIA--LKESELYKSNDKML-- 690
 QY 465 ATTCGSKSRPLNDLVYQKTLQSDITS-----RFSNAIEALNFIQKYSV 510

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 15:11:47 ; Search time 1417.37 Seconds
(without alignments)
1971.845 Million cell updates/sec

Title: US-08-699-716a-1
Perfect score: 1566
Sequence: 1 ATGGCCATCATCATCATCA.....ATGACACGTCGTGTAATCA 1566

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 972840 seqs, 892348106 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl:*
1: gb_da1:*
2: gb_da2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_p11:*
8: gb_p12:*
9: gb_p13:*
10: gb_p14:*
11: gb_p15:*
12: gb_p16:*
13: gb_p17:*
14: gb_p18:*
15: gb_p19:*
16: gb_p20:*
17: gb_p21:*
18: gb_p22:*
19: gb_p23:*
20: gb_p24:*
21: gb_p25:*
22: gb_p26:*
23: gb_p27:*
24: gb_p28:*
25: gb_p29:*
26: gb_p30:*
27: gb_p31:*
28: gb_p32:*
29: gb_p33:*
30: gb_p34:*
31: gb_p35:*
32: gb_p36:*
33: gb_p37:*
34: gb_p38:*
35: gb_p39:*
36: gb_p40:*
37: gb_p41:*
38: gb_p42:*
39: gb_p43:*
40: gb_p44:*
41: gb_p45:*
42: gb_p46:*
43: gb_p47:*

Result No.	Score	Query Match length	DB ID	Description
1	930	59.4	2100	1 YEPICR
2	930	59.4	70305	2 YPCD1
3	930	59.4	70504	2 AF053946
4	930	59.4	70559	2 AF074612
5	927	59.2	1014	5 A46411
6	927	59.2	1014	5 A56793
7	927	59.2	1462	5 A56808
8	927	59.2	1530	5 A56814
9	923	58.9	1014	5 A56795
10	923	58.6	975	2 AF167309
11	918	58.6	981	2 AF167310
12	777	49.6	981	2 AF167310
13	513	32.8	547	5 A56812
14	512	32.7	5383	1 YPCAF
15	512	32.7	96210	2 YPPMT1
16	512	32.7	100984	2 AF053947
17	512	32.7	100990	2 AF074611
18	438	28.0	981	1 YPTPVANT
19	392	25.0	2201	1 YEPICRGVHP
20	259	16.5	1002	1 YEPICRGVHP
21	215	13.7	1002	1 YEPICRGVHP
22	196	12.5	975	1 YEPICRGVHP
23	196	12.5	975	1 YEPICRGVHP
24	189	12.1	975	1 YEPICRGVHP

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

C 171	1.9	111222	49	AC007764	Genomic s	244	30	1.9	177292	68	AC010768	AC010768 Homo sapi
C 172	1.9	111368	9	AP000025	Homo sapi	245	30	1.9	180324	67	AC026212	AC026212 Homo sapi
C 173	1.9	117767	40	AL150262	Homo sapi	246	30	1.9	181098	11	AC004216	AC004216 Homo sapi
C 174	1.9	119118	10	HSRFR7C01	Homo sapi	247	30	1.9	181191	69	AC022135	AC022135 Homo sapi
C 175	1.9	120088	39	AC004865	Homo sapi	248	30	1.9	181727	67	AC026291	AC026291 Homo sapi
C 176	1.9	120937	55	AC010239	Homo sapi	249	30	1.9	183007	54	AC021540	AC021540 Homo sapi
C 177	1.9	121695	32	AL137127	Homo sapi	250	30	1.9	183372	73	AC009161	AC009161 Homo sapi
C 178	1.9	121990	51	AC022314	Arabidops	251	30	1.9	183433	60	AC021907	AC021907 Homo sapi
C 179	1.9	122264	32	HSAL31498	Homo sapi	252	30	1.9	185155	40	AL162414	AL162414 Homo sapi
C 180	1.9	123762	42	AF207954	Homo sapi	253	30	1.9	186206	43	AC009709	AC009709 Homo sapi
C 181	1.9	128293	67	AC016564	Homo sapi	254	30	1.9	187067	42	AC015468	AC015468 Homo sapi
C 182	1.9	128577	67	AC025451	Homo sapi	255	30	1.9	187240	51	AC010066	AC010066 Homo sapi
C 183	1.9	129661	52	AC018543	Homo sapi	256	30	1.9	187956	72	AC021842	AC021842 Homo sapi
C 184	1.9	131007	10	HS998C11	Human DNA	257	30	1.9	188950	77	AC058795	AC058795 Homo sapi
C 185	1.9	131819	11	HS529N6	Human DNA	258	30	1.9	190626	79	AC062036	AC062036 Homo sapi
C 186	1.9	131899	43	AC008258	Drosophila	259	30	1.9	191479	51	AC023173	AC023173 Mus muscu
C 187	1.9	132723	61	AC009848	Drosophila	260	30	1.9	191562	40	AL139350	AL139350 Homo sapi
C 188	1.9	134019	60	AC021361	Homo sapi	261	30	1.9	192751	40	AL158159	AL158159 Homo sapi
C 189	1.9	134593	67	AC008135	Drosophila	262	30	1.9	194275	67	AC025004	AC025004 Homo sapi
C 190	1.9	137938	60	AC011653	Homo sapi	263	30	1.9	194678	71	AC026225	AC026225 Homo sapi
C 191	1.9	138828	43	AC016468	Homo sapi	264	30	1.9	197030	41	AC007914	AC007914 Homo sapi
C 192	1.9	139498	73	AC025297	Homo sapi	265	30	1.9	198310	58	AE003760	AE003760 Drosophila
C 193	1.9	141762	10	HS198C21	Human DNA	266	30	1.9	198320	8	ATCHRIV65	ATCHRIV65 Homo sapi
C 194	1.9	145573	32	AL138904	Homo sapi	267	30	1.9	198549	41	AC009303	AC009303 Homo sapi
C 195	1.9	147157	10	HS545L17	Human DNA	268	30	1.9	208412	43	AC015557	AC015557 Homo sapi
C 196	1.9	149441	55	AC010432	Homo sapi	269	30	1.9	209645	54	AC015575	AC015575 Homo sapi
C 197	1.9	149560	43	AC016050	Homo sapi	270	30	1.9	209826	74	AC021130	AC021130 Homo sapi
C 198	1.9	150204	41	AF16190	Homo sapi	271	30	1.9	212176	39	AC007543	AC007543 Homo sapi
C 199	1.9	150589	78	AC021360	Homo sapi	272	30	1.9	220210	40	AL136360	AL136360 Homo sapi
C 200	1.9	152192	40	AL158814	Homo sapi	273	30	1.9	222302	32	CNSO1DW3	AL136327 Homo sapi
C 201	1.9	153987	80	AC066616	Homo sapi	274	30	1.9	224840	67	AC008763	AC008763 Homo sapi
C 202	1.9	154448	71	AC021391	Homo sapi	275	30	1.9	228508	79	AC063967	AC063967 Mus muscu
C 203	1.9	155987	80	AC066616	Homo sapi	276	30	1.9	228508	79	AC063967	AC063967 Mus muscu
C 204	1.9	159397	80	AC066616	Homo sapi	277	30	1.9	229106	32	AL139331	AL139331 Homo sapi
C 205	1.9	163439	70	AC026508	Homo sapi	278	30	1.9	229106	44	AC010147	AC010147 Homo sapi
C 206	1.9	165182	71	AC019031	Homo sapi	279	30	1.9	229106	32	AL158079	AL158079 Homo sapi
C 207	1.9	167642	54	AC022865	Homo sapi	280	30	1.9	231112	33	PFMALP2	AL158079 Plesmodiu
C 208	1.9	168184	53	AC024158	Homo sapi	281	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 209	1.9	168499	57	AC011642	Homo sapi	282	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 210	1.9	168698	72	AC011070	Drosophila	283	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 211	1.9	168698	72	AC011070	Drosophila	284	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 212	1.9	168698	72	AC011070	Drosophila	285	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 213	1.9	168698	72	AC011070	Drosophila	286	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 214	1.9	168698	72	AC011070	Drosophila	287	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 215	1.9	168698	72	AC011070	Drosophila	288	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 216	1.9	168698	72	AC011070	Drosophila	289	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 217	1.9	168698	72	AC011070	Drosophila	290	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 218	1.9	168698	72	AC011070	Drosophila	291	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 219	1.9	168698	72	AC011070	Drosophila	292	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 220	1.9	168698	72	AC011070	Drosophila	293	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 221	1.9	168698	72	AC011070	Drosophila	294	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 222	1.9	168698	72	AC011070	Drosophila	295	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 223	1.9	168698	72	AC011070	Drosophila	296	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 224	1.9	168698	72	AC011070	Drosophila	297	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 225	1.9	168698	72	AC011070	Drosophila	298	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 226	1.9	168698	72	AC011070	Drosophila	299	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 227	1.9	168698	72	AC011070	Drosophila	300	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 228	1.9	168698	72	AC011070	Drosophila	301	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 229	1.9	168698	72	AC011070	Drosophila	302	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 230	1.9	168698	72	AC011070	Drosophila	303	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 231	1.9	168698	72	AC011070	Drosophila	304	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 232	1.9	168698	72	AC011070	Drosophila	305	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 233	1.9	168698	72	AC011070	Drosophila	306	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 234	1.9	168698	72	AC011070	Drosophila	307	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 235	1.9	168698	72	AC011070	Drosophila	308	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 236	1.9	168698	72	AC011070	Drosophila	309	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 237	1.9	168698	72	AC011070	Drosophila	310	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 238	1.9	168698	72	AC011070	Drosophila	311	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 239	1.9	168698	72	AC011070	Drosophila	312	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 240	1.9	168698	72	AC011070	Drosophila	313	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 241	1.9	168698	72	AC011070	Drosophila	314	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 242	1.9	168698	72	AC011070	Drosophila	315	30	1.9	232885	42	AL157366	AL157366 Homo sapi
C 243	1.9	168698	72	AC011070	Drosophila	316	30	1.9	232885	42	AL157366	AL157366 Homo sapi

C 317	29	1.9	353	13	G10350	human	STS	C	390	29	1.9	9154	42	AC014305	AC014305 Drosophila	
C 318	29	1.9	357	13	G31855	SPAC1	Notwa	C	391	29	1.9	12029	34	AE001381	AE001381 Plasmodium	
C 319	29	1.9	368	13	G08786	human	STS	C	392	29	1.9	20445	39	AC000023	AC000023 Homo sapi	
C 320	29	1.9	377	13	G27240	human	STS	C	393	29	1.9	21961	33	CEP14P4	CEP14P4 Caenorhab	
C 321	29	1.9	392	34	AF140081	Apis	meil	C	394	29	1.9	22209	12	AF100154	AF100154 Rattus no	
C 322	29	1.9	406	13	G09250	human	STS	C	395	29	1.9	23404	42	AC017160	AC017160 Drosophila	
C 323	29	1.9	410	13	G08916	human	STS	C	396	29	1.9	24197	42	AC015221	AC015221 Drosophila	
C 324	29	1.9	412	34	AF077591	Pennaeus	m	C	397	29	1.9	24942	8	YSCD9819	YSCD9819 Saccharomyc	
C 325	29	1.9	426	34	AF169213	Loligo	pe	C	398	29	1.9	25789	7	SCDNACHV	SCDNACHV X91067 S.cerevisia	
C 326	29	1.9	445	13	G41485	Zebraf	C	399	29	1.9	25937	68	AC026346	AC026346 FRRDGEWE		
C 327	29	1.9	445	13	G45889	Z4822.1	zeb	C	400	29	1.9	27470	4	AB023029	AB023029 CEREK849	
C 328	29	1.9	446	13	G45875	Z4740.1	zeb	C	401	29	1.9	29498	7	AB023029	AB023029 CEREK849	
C 329	29	1.9	453	34	AF169214	Loligo	pe	C	402	29	1.9	29544	33	CEK849	CEK849 AB020747 Arabidops	
C 330	29	1.9	458	34	AF140078	Apis	meil	C	403	29	1.9	29605	7	AB017062	AB017062 Arabidops	
C 331	29	1.9	500	13	CNS010RK	Amphel	C	404	29	1.9	30578	33	CEK06G1	CEK06G1 U41014 Caenorhabd		
C 332	29	1.9	510	13	HDMUT6622	Human	STS	C	405	29	1.9	31205	33	CEK06G1	CEK06G1 AC012876 Drosophila	
C 333	29	1.9	522	12	MMT14111	Mus	musculu	C	406	29	1.9	31527	42	AC012876	AC012876 AC017861 Drosophila	
C 334	29	1.9	543	34	AF169207	Loligo	pe	C	407	29	1.9	33606	42	AC017861	AC017861 AC007046 Homo sapi	
C 335	29	1.9	558	5	E02043	untransl	C	408	29	1.9	34506	7	SC8337	SC8337 AC007046 Homo sapi		
C 336	29	1.9	612	13	AU028405	Rattus	no	C	409	29	1.9	34864	11	U73649	U73649 Human Chrom	
C 337	29	1.9	625	13	G09478	human	STS	C	410	29	1.9	35576	41	AC007046	AC007046 Homo sapi	
C 338	29	1.9	631	13	G08223	human	STS	C	411	29	1.9	36170	41	AC007046	AC007046 Homo sapi	
C 339	29	1.9	638	13	G39764	Z10215	zeb	C	412	29	1.9	37353	33	CER1298	CER1298 AC004501 Homo sapi	
C 340	29	1.9	638	13	G39764	Z10215	zeb	C	413	29	1.9	39387	11	AC004501	AC004501 Homo sapi	
C 341	29	1.9	651	13	G03464	human	STS	C	414	29	1.9	40822	10	HS85F18	HS85F18 AC006187 Homo sapi	
C 342	29	1.9	670	7	RIC012	Rice	retrop	C	415	29	1.9	41917	10	HS85F18	HS85F18 AC006187 Homo sapi	
C 343	29	1.9	724	13	G09445	human	STS	C	416	29	1.9	45410	41	AC006187	AC006187 AC010077 Homo sapi	
C 344	29	1.9	799	13	G40184	Z20576	zeb	C	417	29	1.9	46934	39	AC010077	AC010077 AC013346 Homo sapi	
C 345	29	1.9	830	13	G09414	human	STS	C	418	29	1.9	47302	42	AC013346	AC013346 AC062014 Homo sapi	
C 346	29	1.9	905	8	AF030132	Arabidops	C	419	29	1.9	54368	78	AC062014	AC062014 AC019976 Drosophila		
C 347	29	1.9	924	39	AF164688	Homo sapi	C	420	29	1.9	59762	7	AB023032	AB023032 Arabidops		
C 348	29	1.9	926	13	CNS0110E	Anopheles	C	421	29	1.9	59762	7	AB023032	AB023032 Arabidops		
C 349	29	1.9	1021	81	ORCPV160	Cowpox	viru	C	422	29	1.9	61456	42	AC014553	AC014553 Drosophila	
C 350	29	1.9	1033	81	HAU67261	Helicoverpa	C	423	29	1.9	61756	42	AC014553	AC014553 Drosophila		
C 351	29	1.9	1337	34	AF114166	Cryptospor	C	424	29	1.9	63250	42	AC015767	AC015767 Homo sapi		
C 352	29	1.9	1508	58	S55234	Dictyostell	C	425	29	1.9	64392	78	AC060795	AC060795 Homo sapi		
C 353	29	1.9	1585	11	HSX800565	Dictyostell	C	426	29	1.9	65577	51	AC023689	AC023689 Drosophila		
C 354	29	1.9	1650	11	AF046911	Mus	musculu	C	427	29	1.9	65796	51	AC023570	AC023570 Homo sapi	
C 355	29	1.9	1733	3	AC09772	Amphidini	C	428	29	1.9	66940	45	AC005642	AC005642 Drosophila		
C 356	29	1.9	1792	33	S76003	lens-specific	C	429	29	1.9	68053	42	AC015506	AC015506 Homo sapi		
C 357	29	1.9	1792	34	S97760	S-crySTALL	C	430	29	1.9	70447	57	AC025722	AC025722 Caenorhab		
C 358	29	1.9	1813	12	MUSIGHV2	Mouse	Ig	C	431	29	1.9	71345	75	AC044830	AC044830 Homo sapi	
C 359	29	1.9	1888	7	ATH0109F	Arabidops	C	432	29	1.9	71807	7	AB025613	AB025613 Arabidops		
C 360	29	1.9	1920	33	OMMCRYSB1	M74321	O.sloanei	C	433	29	1.9	72068	76	AC055785	AC055785 Homo sapi	
C 361	29	1.9	1936	33	DMV18	X15382	D.alscoiden	C	434	29	1.9	72456	71	AC027728	AC027728 Homo sapi	
C 362	29	1.9	2315	11	HSX800321	M74041	I.punctatus	C	435	29	1.9	72659	52	AC010174	AC010174 Homo sapi	
C 363	29	1.9	2316	4	ICRIGH	A37615	Sequence 1	C	436	29	1.9	74516	57	AC025641	AC025641 Homo sapi	
C 364	29	1.9	2359	5	A37615	Z11786	C.parvum	C	437	29	1.9	74877	7	AB006705	AB006705 Arabidops	
C 365	29	1.9	2359	33	CPWALPR	M95743	Cryptospor	C	438	29	1.9	75596	72	AC034122	AC034122 Mus musculu	
C 366	29	1.9	2359	33	CYDINA	M60052	Human histi	C	439	29	1.9	77834	43	AC021610	AC021610 Homo sapi	
C 367	29	1.9	2365	10	HDMHCPB	AF124454	Schizocoe	C	440	29	1.9	79469	49	ATAC010852	ATAC010852 Arabidops	
C 368	29	1.9	2392	34	AF124454	Y14756	Mus musculu	C	441	29	1.9	79597	39	AF157623	AF157623 Homo sapi	
C 369	29	1.9	2475	12	MMSTOEX4	M95580	Yeast ORF1	C	442	29	1.9	80417	70	AC027562	AC027562 Homo sapi	
C 370	29	1.9	2483	7	YSCSEOCAT	U63062	Dictyostell	C	443	29	1.9	80442	49	T24D18	T24D18 AC010924 Arabidops	
C 371	29	1.9	2720	33	DDU63062	AF054907	Neutrospor	C	444	29	1.9	80970	7	AB006697	AB006697 Arabidops	
C 372	29	1.9	2725	8	AF054907	U17641	Neutrospor	C	445	29	1.9	81542	7	AB026647	AB026647 Arabidops	
C 373	29	1.9	2855	7	NC017641	AF006304	Saccharom	C	446	29	1.9	82415	7	AB005244	AB005244 Arabidops	
C 374	29	1.9	3069	8	AF006304	AF081801	Dictyoste	C	447	29	1.9	82571	55	AC025116	AC025116 Mus musculu	
C 375	29	1.9	3386	7	ACA9670	AF109776	Metapeneae	C	448	29	1.9	82603	43	AC021939	AC021939 Arabidops	
C 376	29	1.9	3625	34	AF081801	Z36019	S.cerevisia	C	449	29	1.9	83097	7	AB025624	AB025624 Arabidops	
C 377	29	1.9	3705	34	AF109776	Z74796	S.cerevisia	C	450	29	1.9	83535	69	AC012619	AC012619 Homo sapi	
C 378	29	1.9	4009	7	SCYBR150C	J05161	Earthworm	C	451	29	1.9	83544	7	AB025615	AB025615 Arabidops	
C 379	29	1.9	4037	33	LUMHBC	M61742	A.thaliana	C	452	29	1.9	83957	10	HS885517	HS885517 Arabidops	
C 380	29	1.9	4037	33	LUMHBC	Z74797	S.cerevisia	C	453	29	1.9	85992	7	AB009052	AB009052 Arabidops	
C 381	29	1.9	4418	7	ATHPATPC	D00319	Cowpox	C	454	29	1.9	86659	53	AC020665	AC020665 Arabidops	
C 382	29	1.9	4434	7	SCYLOJ55C	U23477	Dictyostell	C	455	29	1.9	86771	3	AC007190	AC007190 Genomic s	
C 383	29	1.9	4644	81	CPVATI	Z22537	C.parvum	C	456	29	1.9	87187	7	AB015474	AB015474 Arabidops	
C 384	29	1.9	5081	43	AC019721	M25161	Human Na,K-	C	457	29	1.9	87967	8	AC005223	AC005223 Arabidops	
C 385	29	1.9	5577	33	DDU23477	U18814	Saccharomyc	C	458	29	1.9	88031	51	AC004378	AC004378 Drosophila	
C 386	29	1.9	6030	33	CPOCYWLP4	AC014339	Drosophila	C	459	29	1.9	89338	49	AC006068	AC006068 Arabidops	
C 387	29	1.9	7675	9	HDMNAKATP1			C	460	29						
C 388	29	1.9	8443	8	SCE3612			C	461	29						
C 389	29	1.9	8661	42	AC014329			C	462	29						

463	29	1.9	89779	7	AB005234	AB005234 Arabidops	C 536	29	1.9	131547	72	AC011380	AC011380 Homo sapi
464	29	1.9	90065	34	AC005147	AC005147 Drosophill	C 537	29	1.9	132596	42	AC013046	AC013046 Drosophill
465	29	1.9	91480	10	HS04823F3	AL079333 Homo sapi	C 538	29	1.9	133318	41	AC007770	AC007770 Drosophill
466	29	1.9	92153	41	AL353802	AL353802 Homo sapi	C 539	29	1.9	133774	53	AC016007	AC016007 Homo sapi
467	29	1.9	94972	49	AC005700	AC005700 Arabidops	C 540	29	1.9	134292	43	AC019051	AC019051 Homo sapi
468	29	1.9	95695	80	AC051631	AC051631 Arabidops	C 541	29	1.9	135038	11	H004703	H004703 Homo sapi
469	29	1.9	95933	43	AC015848	AC015848 Homo sapi	C 542	29	1.9	135149	55	AC018818	AC018818 Homo sapi
470	29	1.9	96075	42	AC018294	AC018294 Homo sapi	C 543	29	1.9	135777	56	AC024510	AC024510 Homo sapi
471	29	1.9	96902	39	AC012599	AC012599 Homo sapi	C 544	29	1.9	136842	76	AC027118	AC027118 Homo sapi
472	29	1.9	97263	49	ATAC013258	AC013258 Arabidops	C 545	29	1.9	137438	67	AC024993	AC024993 Homo sapi
473	29	1.9	98274	78	AC062019	AC062019 Homo sapi	C 546	29	1.9	137535	55	AC020914	AC020914 Homo sapi
474	29	1.9	99332	11	HS079818	AL096764 Homo sapi	C 547	29	1.9	137650	41	AF161801	AF161801 Homo sapi
475	29	1.9	99357	72	AC020933	AC020933 Homo sapi	C 548	29	1.9	137938	60	AC011653	AC011653 Homo sapi
476	29	1.9	100000	9	AB020877	AB020877 Homo sapi	C 549	29	1.9	138187	12	MM004461	MM004461 Homo sapi
477	29	1.9	100000	9	AP000069	AP000069 Homo sapi	C 550	29	1.9	139480	39	H00400331	H00400331 Homo sapi
478	29	1.9	100000	9	AP000499	AP000499 Homo sapi	C 551	29	1.9	139820	39	AC005100	AC005100 Homo sapi
479	29	1.9	100364	39	AC007590	AC007590 Homo sapi	C 552	29	1.9	140136	42	AC017381	AC017381 Homo sapi
480	29	1.9	100835	49	ATAC016795	AC016795 Arabidops	C 553	29	1.9	140644	70	AC016797	AC016797 Homo sapi
481	29	1.9	100975	8	AC007980	AC007980 Arabidops	C 554	29	1.9	141056	40	HS1184F4	HS1184F4 Homo sapi
482	29	1.9	101371	8	AC002534	AC002534 Arabidops	C 555	29	1.9	141201	60	AC022637	AC022637 Homo sapi
483	29	1.9	102621	33	DMBR43E12	AL138971 Drosophill	C 556	29	1.9	141243	57	AC008286	AC008286 Homo sapi
484	29	1.9	103225	55	AC022405	AC022405 Homo sapi	C 557	29	1.9	141391	72	AC024309	AC024309 Homo sapi
485	29	1.9	104607	8	ATF28J14	AL163652 Arabidops	C 558	29	1.9	142651	67	HS202121	HS202121 Homo sapi
486	29	1.9	105306	8	ATF9D24	AL137081 Arabidops	C 559	29	1.9	142665	10	HS33B19	HS33B19 Homo sapi
487	29	1.9	105448	50	AC010494	AC010494 Homo sapi	C 560	29	1.9	142807	42	AC007812	AC007812 Homo sapi
488	29	1.9	105733	49	AC012561	AC012561 Arabidops	C 561	29	1.9	143240	11	HS309H15	HS309H15 Homo sapi
489	29	1.9	105937	49	ATAC012679	AC012679 Arabidops	C 562	29	1.9	143429	42	AC017124	AC017124 Homo sapi
490	29	1.9	106482	12	AC007585	AC007585 Homo sapi	C 563	29	1.9	143450	78	AC061711	AC061711 Homo sapi
491	29	1.9	106870	42	AF212852	AF212852 Homo sapi	C 564	29	1.9	143459	48	HS044211	HS044211 Homo sapi
492	29	1.9	106936	42	AC009909	AC009909 Drosophill	C 565	29	1.9	144639	32	HS044211	HS044211 Homo sapi
493	29	1.9	107950	7	ATT17F15	AL049658 Arabidops	C 566	29	1.9	144784	56	AC015761	AC015761 Homo sapi
494	29	1.9	109016	8	ATT10K17	AL132977 Arabidops	C 567	29	1.9	145481	53	AC009338	AC009338 Homo sapi
495	29	1.9	110000	31	CEV63D3_1	Continuation (2 of	C 568	29	1.9	145538	10	HS245G19	HS245G19 Homo sapi
496	29	1.9	110000	32	CEV105E8_0	AL022594 Ctenothab	C 569	29	1.9	145530	70	AC018572	AC018572 Homo sapi
497	29	1.9	110000	32	HS1214H10_0	AL022344 Homo sapi	C 570	29	1.9	145835	59	AC020560	AC020560 Homo sapi
498	29	1.9	110000	32	HS1214H10_1	Continuation (2 of	C 571	29	1.9	145870	56	AC023581	AC023581 Homo sapi
499	29	1.9	111524	42	AF206725	AF206725 Homo sapi	C 572	29	1.9	146468	69	AC010765	AC010765 Homo sapi
500	29	1.9	111768	10	HS879J18	AL035400 Homo sapi	C 573	29	1.9	146501	62	AC022433	AC022433 Homo sapi
501	29	1.9	111945	8	ATF1C12	AL022324 Arabidops	C 574	29	1.9	147765	55	AC010412	AC010412 Homo sapi
502	29	1.9	112206	51	AC008317	AC008317 Drosophill	C 575	29	1.9	147812	78	AC062025	AC062025 Homo sapi
503	29	1.9	112460	10	HS1104E15	AL022317 Homo sapi	C 576	29	1.9	147902	11	AC005670	AC005670 Homo sapi
504	29	1.9	112930	55	AC008795	AL035704 Homo sapi	C 577	29	1.9	148606	11	HSB451C14	HSB451C14 Homo sapi
505	29	1.9	113956	40	HS74M1	AL033704 Homo sapi	C 578	29	1.9	148640	4	AF056116	AF056116 Homo sapi
506	29	1.9	114958	51	AC008356	AC008356 Drosophill	C 579	29	1.9	149337	76	AC034150	AC034150 Homo sapi
507	29	1.9	116177	8	T13D8	AC004473 Arabidops	C 580	29	1.9	149447	40	AL353139	AL353139 Homo sapi
508	29	1.9	116218	39	AC005090	AC005090 Homo sapi	C 581	29	1.9	149594	31	AP000856	AP000856 Homo sapi
509	29	1.9	116370	10	HS569M23	AL031666 Homo sapi	C 582	29	1.9	149748	61	AC021236	AC021236 Homo sapi
510	29	1.9	117406	50	AC010792	AC010792 Homo sapi	C 583	29	1.9	150445	32	AL161742	AL161742 Homo sapi
511	29	1.9	118322	55	AC008891	AC008891 Homo sapi	C 584	29	1.9	150482	31	AP001388	AP001388 Homo sapi
512	29	1.9	118595	39	AC002379	AC002379 Homo sapi	C 585	29	1.9	151429	11	AC005279	AC005279 Homo sapi
513	29	1.9	118893	32	AL138839	AL138839 Homo sapi	C 586	29	1.9	151670	56	AC022669	AC022669 Homo sapi
514	29	1.9	118931	41	AF188031	AF188031 Homo sapi	C 587	29	1.9	152512	72	AC025799	AC025799 Homo sapi
515	29	1.9	121290	77	AC034256	AC034256 Arabidops	C 588	29	1.9	152544	79	AC011309	AC011309 Homo sapi
516	29	1.9	122796	73	AC009135	AC009135 Homo sapi	C 589	29	1.9	152604	78	AC016671	AC016671 Homo sapi
517	29	1.9	122877	53	AC015768	AC015768 Homo sapi	C 590	29	1.9	153009	41	AC009801	AC009801 Homo sapi
518	29	1.9	123183	49	AC006841	AC006841 Arabidops	C 591	29	1.9	153098	33	PFA13P2	PFA13P2 Homo sapi
519	29	1.9	123695	11	HS058113	AL138826 Homo sapi	C 592	29	1.9	153205	60	AC025828	AC025828 Homo sapi
520	29	1.9	123766	41	AC007813	AC007813 Drosophill	C 593	29	1.9	153234	78	AC026947	AC026947 Homo sapi
521	29	1.9	124700	34	AC000782	AC000782 Drosophill	C 594	29	1.9	153323	77	AC023665	AC023665 Homo sapi
522	29	1.9	124700	34	AC000558	AC000558 Drosophill	C 595	29	1.9	153704	69	AC020919	AC020919 Homo sapi
523	29	1.9	125644	73	AC034234	AL157403 Homo sapi	C 596	29	1.9	153738	74	AC040161	AC040161 Homo sapi
524	29	1.9	125900	32	AL157403	AL157403 Homo sapi	C 597	29	1.9	154103	76	AC051622	AC051622 Homo sapi
525	29	1.9	127587	41	AC005959	AC005959 Homo sapi	C 598	29	1.9	154788	10	HS325F22	HS325F22 Homo sapi
526	29	1.9	127735	10	HS257E24	AL034424 Homo sapi	C 599	29	1.9	155021	39	AC007159	AC007159 Homo sapi
527	29	1.9	127867	10	HS694B14	AL031673 Homo sapi	C 600	29	1.9	155069	78	AC018730	AC018730 Homo sapi
528	29	1.9	128117	39	AC004000	AC004000 Homo sapi	C 601	29	1.9	155125	52	AC021069	AC021069 Homo sapi
529	29	1.9	128132	55	AC007983	AC007983 Drosophill	C 602	29	1.9	155323	68	AC009824	AC009824 Homo sapi
530	29	1.9	128722	10	HS15005	AL021154 Homo sapi	C 603	29	1.9	155450	11	AC005951	AC005951 Homo sapi
531	29	1.9	129659	10	HS204E5	Z98941 Homo sapi	C 604	29	1.9	155526	57	AC013371	AC013371 Homo sapi
532	29	1.9	130472	32	AL188041	AL188041 Homo sapi	C 605	29	1.9	155537	59	AC013072	AC013072 Homo sapi
533	29	1.9	130668	11	AC005736	AC005736 Homo sapi	C 606	29	1.9	155795	70	AC027270	AC027270 Homo sapi
534	29	1.9	130755	80	AC004253	AL031280 Homo sapi	C 607	29	1.9	155829	79	AC013710	AC013710 Homo sapi
535	29	1.9	131318	10	HS187B23	AL031280 Homo sapi	C 608	29	1.9	155998	54	AC015567	AC015567 Homo sapi

609	29	1.9 156843	79	AC011145	Homo sapi	682	29	1.9 171772	42	AC009584	Homo sapi
610	29	1.9 156928	75	AC026937	Homo sapi	683	29	1.9 172104	57	AC024933	Homo sapi
611	29	1.9 157289	50	AC009234	Homo sapi	684	29	1.9 172913	53	AC023291	Homo sapi
612	29	1.9 157320	55	AC011365	Homo sapi	685	29	1.9 173116	31	AP001375	Homo sapi
613	29	1.9 157321	39	AC004912	Homo sapi	686	29	1.9 173309	61	AC025037	Homo sapi
614	29	1.9 157340	73	AC015652	Homo sapi	687	29	1.9 173781	34	AC005652	Homo sapi
615	29	1.9 157587	32	AL139125	Homo sapi	688	29	1.9 173932	71	AC013793	Homo sapi
616	29	1.9 157751	41	AC007098	Homo sapi	689	29	1.9 174090	54	AC021330	Homo sapi
617	29	1.9 157776	32	AL138893	Homo sapi	690	29	1.9 174517	76	AC009927	Homo sapi
618	29	1.9 158213	73	AC015669	Homo sapi	691	29	1.9 174731	42	AC012703	Homo sapi
619	29	1.9 158441	60	AC016803	Homo sapi	692	29	1.9 174732	42	AC012160	Homo sapi
620	29	1.9 158623	54	AC010921	Homo sapi	693	29	1.9 174977	69	AC022608	Homo sapi
621	29	1.9 158889	42	AC008031	Typanoso	694	29	1.9 175167	72	AC011404	Homo sapi
622	29	1.9 159270	41	AC007241	Homo sapi	695	29	1.9 175441	54	AC022869	Homo sapi
623	29	1.9 159393	31	AP001330	Homo sapi	696	29	1.9 175516	60	AC006280	Plasmodu
624	29	1.9 161136	32	AL138721	Homo sapi	697	29	1.9 175657	74	AC027663	Homo sapi
625	29	1.9 161739	78	AC018868	Homo sapi	698	29	1.9 176140	77	AC034172	Homo sapi
626	29	1.9 161778	56	AC024892	Homo sapi	699	29	1.9 176186	39	AC006825	Homo sapi
627	29	1.9 161930	75	AC026828	Homo sapi	700	29	1.9 176327	75	AC018859	Homo sapi
628	29	1.9 161955	77	AC034100	Mus muscu	701	29	1.9 176424	31	AP000923	Homo sapi
629	29	1.9 162013	11	CNS01DMQ	AL137918	702	29	1.9 176595	40	CNS05TBN	AL163872
630	29	1.9 162481	8	AF061282	Sorghum b	703	29	1.9 176612	57	AC021308	Homo sapi
631	29	1.9 162481	81	AF061282	Sorghum b	704	29	1.9 176895	55	AC010327	Homo sapi
632	29	1.9 162532	78	AC021014	Homo sapi	705	29	1.9 177802	59	AC018718	Homo sapi
633	29	1.9 162636	43	AC009551	Homo sapi	706	29	1.9 178181	41	AL135368	Homo sapi
634	29	1.9 162822	11	AC005751	Homo sapi	707	29	1.9 178477	61	AC022801	Homo sapi
635	29	1.9 162925	70	AC018528	Homo sapi	708	29	1.9 178683	40	AL136129	Homo sapi
636	29	1.9 163062	71	AC022352	Oryza sat	709	29	1.9 179798	79	AC015978	Homo sapi
637	29	1.9 163277	76	AC010095	Homo sapi	710	29	1.9 179880	69	AC011679	Homo sapi
638	29	1.9 163373	74	AC024263	Homo sapi	711	29	1.9 179884	40	AL161896	Homo sapi
639	29	1.9 163420	56	AC012576	Homo sapi	712	29	1.9 180046	42	AL161896	Homo sapi
640	29	1.9 163779	68	AC026670	Homo sapi	713	29	1.9 180112	42	AC013464	Homo sapi
641	29	1.9 163780	75	AC025806	Homo sapi	714	29	1.9 180385	68	AC021884	Homo sapi
642	29	1.9 163880	41	AC007597	Homo sapi	715	29	1.9 180485	69	AC027145	Homo sapi
643	29	1.9 163907	41	AC011114	Homo sapi	716	29	1.9 180737	70	AC026472	Homo sapi
644	29	1.9 164139	44	AC012593	Homo sapi	717	29	1.9 180780	44	AC010683	Homo sapi
645	29	1.9 164399	33	PF041236	Z98551	718	29	1.9 180919	39	AF088219	Homo sapi
646	29	1.9 164437	32	AL137848	Homo sapi	719	29	1.9 181060	61	AC021037	Homo sapi
647	29	1.9 164548	31	CEY70D2	AL008880	720	29	1.9 181150	78	AC017020	Homo sapi
648	29	1.9 164945	61	AC024968	Caenorhab	721	29	1.9 181183	42	AC013288	Homo sapi
649	29	1.9 164945	61	AC024968	Caenorhab	722	29	1.9 181313	61	AC021963	Homo sapi
650	29	1.9 165000	40	AL138895	Homo sapi	723	29	1.9 181466	79	AC005073	Homo sapi
651	29	1.9 166276	54	AC019239	Homo sapi	724	29	1.9 181541	59	AC013445	Homo sapi
652	29	1.9 166325	42	AC013271	Homo sapi	725	29	1.9 182049	68	AC007923	Homo sapi
653	29	1.9 166335	70	AC011946	Homo sapi	726	29	1.9 182798	69	AC025394	Homo sapi
654	29	1.9 166416	53	AC016866	Homo sapi	727	29	1.9 183082	74	AC010138	Homo sapi
655	29	1.9 166445	78	AC009898	Homo sapi	728	29	1.9 183210	40	AL162742	Homo sapi
656	29	1.9 166701	11	AC005747	Homo sapi	729	29	1.9 183575	40	AL161648	Homo sapi
657	29	1.9 166807	78	AC008060	Homo sapi	730	29	1.9 183837	57	AC022613	Homo sapi
658	29	1.9 166832	41	AC011830	Homo sapi	731	29	1.9 183855	56	AC022770	Homo sapi
659	29	1.9 166892	39	AC007463	Homo sapi	732	29	1.9 183963	69	AC022430	Homo sapi
660	29	1.9 166942	10	HS1054A22	AL031651	733	29	1.9 184031	60	AC020782	Homo sapi
661	29	1.9 166944	40	AL353072	Human DNA	734	29	1.9 184040	71	AC011877	Homo sapi
662	29	1.9 167082	53	AC021472	Homo sapi	735	29	1.9 184399	53	AC020573	Homo sapi
663	29	1.9 167722	44	AC021243	Homo sapi	736	29	1.9 184460	71	AC027677	Homo sapi
664	29	1.9 167825	11	CNS01BOU	AL109662	737	29	1.9 184505	32	AL139190	Homo sapi
665	29	1.9 168468	39	AC005510	Human chr	738	29	1.9 184921	54	AC016297	Homo sapi
666	29	1.9 168481	67	AC025094	Homo sapi	739	29	1.9 184949	39	AC005029	Homo sapi
667	29	1.9 168481	43	AC015711	Homo sapi	740	29	1.9 185019	54	AC002131	Homo sapi
668	29	1.9 168620	61	AC021732	Homo sapi	741	29	1.9 185249	11	CNS01DTM	AL132800
669	29	1.9 168837	79	AC007629	Homo sapi	742	29	1.9 185461	31	AP001107	Homo sapi
670	29	1.9 169287	40	AL157781	Homo sapi	743	29	1.9 185471	59	AC019208	Homo sapi
671	29	1.9 169308	78	AC041047	Homo sapi	744	29	1.9 186087	72	AC026003	Homo sapi
672	29	1.9 169443	32	AL139809	Homo sapi	745	29	1.9 186212	40	AL161616	Homo sapi
673	29	1.9 170046	40	AL162586	Homo sapi	746	29	1.9 187150	72	AC025924	Homo sapi
674	29	1.9 170561	67	AC008690	Homo sapi	747	29	1.9 187345	54	AC011076	Homo sapi
675	29	1.9 170894	79	AC007404	Homo sapi	748	29	1.9 187396	32	AL157758	Homo sapi
676	29	1.9 170896	70	AC011010	Homo sapi	749	29	1.9 187883	39	AC008268	Homo sapi
677	29	1.9 170970	73	AC021564	Homo sapi	750	29	1.9 188080	42	AC012195	Homo sapi
678	29	1.9 171005	55	AC018495	Homo sapi	751	29	1.9 188735	11	CNS00MBM	AL079305
679	29	1.9 171057	44	AC012470	Homo sapi	752	29	1.9 189625	63	AC021547	Human chr
680	29	1.9 171398	60	AC022060	Homo sapi	753	29	1.9 190159	40	AL137248	Homo sapi
681	29	1.9 171532	78	AC016993	Homo sapi	754	29	1.9 190650	60	AC013799	Homo sapi

755	29	1.9 191060	60	AC021468	Homo sapi	828	29	1.9 259920	58	AE003828	AE003828 Drosophila
756	29	1.9 192366	34	AC008369	Drosophila	829	29	1.9 260067	58	AE003807	AE003807 Drosophila
757	29	1.9 192526	68	AC021138	Homo sapi	830	29	1.9 260104	34	AE003687	AE003687 Drosophila
758	29	1.9 192590	39	AC009509	Homo sapi	831	29	1.9 260104	34	AE003687	AE003687 Drosophila
759	29	1.9 192791	42	AC009364	Homo sapi	832	29	1.9 264996	40	AL137016	AL137016 Homo sapi
760	29	1.9 192828	78	AC017036	Homo sapi	833	29	1.9 264996	40	AL137016	AL137016 Homo sapi
761	29	1.9 193224	78	AC025668	Mus muscu	834	29	1.9 265562	70	AC027308	AC027308 Homo sapi
762	29	1.9 193224	78	AC025668	Mus muscu	835	29	1.9 267770	53	AC016745	AC016745 Homo sapi
763	29	1.9 193881	32	AL157386	Homo sapi	836	29	1.9 268100	42	AC018643	AC018643 Homo sapi
764	29	1.9 193930	79	AC017111	Homo sapi	837	29	1.9 269082	31	CEY7642	CEY7642 Homo sapi
765	29	1.9 193981	76	AC016826	Homo sapi	838	29	1.9 269542	53	AC021155	AC021155 Homo sapi
766	29	1.9 194137	31	AP001446	Homo sapi	839	29	1.9 269641	54	AC008482	AC008482 Homo sapi
767	29	1.9 194317	40	AL139824	Homo sapi	840	29	1.9 270000	10	AB026898	AB026898 Homo sapi
768	29	1.9 194562	68	AC025408	Homo sapi	841	29	1.9 272476	60	AC025809	AC025809 Homo sapi
769	29	1.9 194573	44	AC016914	Homo sapi	842	29	1.9 275185	54	AC008478	AC008478 Homo sapi
770	29	1.9 194841	71	AC021139	Homo sapi	843	29	1.9 281502	34	AE003546	AE003546 Drosophila
771	29	1.9 194905	68	AC009517	Homo sapi	844	29	1.9 289755	51	AC020664	AC020664 Homo sapi
772	29	1.9 195470	43	AC022259	Homo sapi	845	29	1.9 295312	34	AE003582	AE003582 Homo sapi
773	29	1.9 195653	39	AC005873	ctb_15_o	846	29	1.9 298336	34	AE003504	AE003504 Homo sapi
774	29	1.9 196993	32	AL158829	Homo sapi	847	29	1.9 299012	34	AE003570	AE003570 Homo sapi
775	29	1.9 197045	76	AC016821	Homo sapi	848	29	1.9 300205	34	AE003425	AE003425 Homo sapi
776	29	1.9 197071	52	AC020788	Homo sapi	849	29	1.9 303648	34	AE003486	AE003486 Homo sapi
777	29	1.9 197357	70	AC019049	Homo sapi	850	29	1.9 311662	32	AL136181	AL136181 Homo sapi
778	29	1.9 197540	72	AC023555	Homo sapi	851	29	1.9 325686	43	AC021094	AC021094 Homo sapi
779	29	1.9 197791	32	AL136178	Homo sapi	852	29	1.9 329391	34	AE003681	AE003681 Homo sapi
780	29	1.9 198427	8	ATCHRIV52	Arabidops	853	29	1.9 335282	53	AC017085	AC017085 Homo sapi
781	29	1.9 198948	8	AC005675	Drosophila	854	28	1.8 76	15	AX002271	AX002271 Sequence
782	29	1.9 199069	32	AL136380	Homo sapi	855	28	1.8 118	14	SCPI04123	SCPI04123 Artificial
783	29	1.9 199612	55	AC016543	Homo sapi	856	28	1.8 161	8	AF027616	AF027616 Arabidops
784	29	1.9 199957	44	AC012043	Homo sapi	857	28	1.8 260	12	RN089598	RN089598 Rattus norv
785	29	1.9 200746	40	AL133343	Homo sapi	858	28	1.8 277	13	G09999	G09999 human STS C
786	29	1.9 201990	39	AC002385	Human BAC	859	28	1.8 301	13	G10326	G10326 human STS C
787	29	1.9 202192	55	AC023762	Homo sapi	860	28	1.8 344	13	HM071995	HM071995 human STS U
788	29	1.9 202223	11	AC003664	Homo sapi	861	28	1.8 355	13	G47862	G47862 human STS U
789	29	1.9 202551	31	AP001484	Homo sapi	862	28	1.8 410	13	HM078001	HM078001 human STS U
790	29	1.9 203226	43	AC017002	Homo sapi	863	28	1.8 433	13	AF140072	AF140072 human STS U
791	29	1.9 203228	40	AL135903	Homo sapi	864	28	1.8 527	34	GA0358	GA0358 zebrafish
792	29	1.9 203418	75	AC026011	Homo sapi	865	28	1.8 574	13	AF108931	AF108931 human STS U
793	29	1.9 204617	70	AC025644	Homo sapi	866	28	1.8 706	34	AF108931	AF108931 human STS U
794	29	1.9 204855	32	AL137249	Homo sapi	867	28	1.8 858	13	CNS011VF	CNS011VF human STS U
795	29	1.9 205685	44	AC013446	Homo sapi	868	28	1.8 900	10	CNS011ETN	CNS011ETN human STS U
796	29	1.9 206476	52	AC012557	Homo sapi	869	28	1.8 979	13	CNS011G25	CNS011G25 human STS U
797	29	1.9 207375	32	CNS0048M	Homo sapi	870	28	1.8 986	7	SASCHS1	SASCHS1 human STS U
798	29	1.9 207608	76	AC022397	Homo sapi	871	28	1.8 1004	13	CNS011E6	CNS011E6 human STS U
799	29	1.9 207722	79	AC019206	Homo sapi	872	28	1.8 1036	13	CNS01106	CNS01106 human STS U
800	29	1.9 207869	69	AC023050	Homo sapi	873	28	1.8 1310	15	AX002276	AX002276 Sequence
801	29	1.9 208363	11	AC003101	Homo sapi	874	28	1.8 1675	33	BMRDNTX	BMRDNTX human STS U
802	29	1.9 208673	68	AC026100	Homo sapi	875	28	1.8 2005	33	SCSMOXIT	SCSMOXIT human STS U
803	29	1.9 208949	53	AC024171	Homo sapi	876	28	1.8 2496	34	AF12312	AF12312 human STS U
804	29	1.9 209382	39	AC005071	Homo sapi	877	28	1.8 2873	7	SASCHS2	SASCHS2 human STS U
805	29	1.9 209787	44	AC012493	Homo sapi	878	28	1.8 3000	49	AF163819	AF163819 Arabidops
806	29	1.9 209872	57	AC012301	Homo sapi	879	28	1.8 3263	15	AX002278	AX002278 Sequence
807	29	1.9 210829	44	AC012506	Homo sapi	880	28	1.8 3324	10	HSR270993	HSR270993 human STS U
808	29	1.9 211188	55	AC024093	Homo sapi	881	28	1.8 3904	7	SCYDR080M	SCYDR080M human STS U
809	29	1.9 211892	55	AC008756	Homo sapi	882	28	1.8 4001	34	AF067153	AF067153 human STS U
810	29	1.9 213813	73	AC007610	Homo sapi	883	28	1.8 4273	2	AF012911	AF012911 human STS U
811	29	1.9 215574	40	AL157710	Homo sapi	884	28	1.8 4286	15	AX002277	AX002277 Sequence
812	29	1.9 218094	43	AC015797	Mus muscu	885	28	1.8 4671	5	AS8998	AS8998 Sequence 2
813	29	1.9 218741	73	AC008869	Homo sapi	886	28	1.8 4742	81	AF015297	AF015297 Human her
814	29	1.9 219063	53	AC016943	Homo sapi	887	28	1.8 4965	14	AF147464	AF147464 T7 expires
815	29	1.9 219151	73	AC012175	Homo sapi	888	28	1.8 5789	7	AB000223	AB000223 Yeast DNA
816	29	1.9 223906	40	AL162501	Homo sapi	889	28	1.8 7249	42	AC017839	AC017839 Drosophila
817	29	1.9 226905	58	AE003773	Drosophila	890	28	1.8 8733	34	AF121313	AF121313 Drosophila
818	29	1.9 230081	52	AC009387	Homo sapi	891	28	1.8 11572	34	AF177386	AF177386 Drosophila
819	29	1.9 240257	41	HSBA416N2	Homo sapi	892	28	1.8 11579	34	AF146362	AF146362 Drosophila
820	29	1.9 244951	32	AL139084	Homo sapi	893	28	1.8 12351	42	AC014502	AC014502 Drosophila
821	29	1.9 246856	53	AC013242	Homo sapi	894	28	1.8 12575	81	LPICMDHC	LPICMDHC human STS U
822	29	1.9 248778	41	MAE000663	Mus muscu	895	28	1.8 13965	81	AF015298	AF015298 Human her
823	29	1.9 250611	12	AC009774	Homo sapi	896	28	1.8 19451	42	AC014476	AC014476 Drosophila
824	29	1.9 251384	41	HS1191N16	Homo sapi	897	28	1.8 21191	81	HM013194	HM013194 Human herpe
825	29	1.9 251726	69	AC008737	Homo sapi	898	28	1.8 27621	43	AC020338	AC020338 Drosophila
826	29	1.9 252455	41	AC009260	Homo sapi	899	28	1.8 32821	7	SCCHROIV	SCCHROIV human STS U
827	29	1.9 253747	34	AE003723	Drosophila	900	28	1.8 33455	33	CEM0448	CEM0448 Caenorhabditis


```

CDS
481..1461
/gene="lcrv"
/codon_start=1
/transl_table=11
/protein_id="AAA27641.1"
/db_xref="GI:155450"
/translation="MIRAEONPHQFIEDLEKRVQULTGHSSVLEELVOLYKDKNT
DISIKYDRKSEVFNANVITDDIELKRIAYFLPEDAILKGHYDQLONGIRK
EPLSSPTOMELRAFMVMSHSLADRIDDLKIVDSNMHGDARSKLEELAEI
TAEIKYVIOAEIKNHLSSSGTINIHKRSIMDKNLYGTDEIEFRAAEYKILEK
MPTTIOYDSEKRIVSIKDFISENKRGTAGNKNKNSYKKNDELNSHRTTSDSK
SRPLNDVYSOKTQLSDITSRNSAIEALNRTQIXDYMORLDDTSGK"
RBS
1463..1468
/gene="lcrv"
1465..1980
/gene="lcrv"
1474..1980
/gene="lcrv"
/codon_start=1
/transl_table=11
/protein_id="AAA27642.1"
/db_xref="GI:155451"
/translation="MQQETTDPTQEQVLAEMSEFLKGGCTIAMIENISSDTLESLAT
NTOYSGKEDHRYQALCVLDHDSREFFLGACROAMGODLAIHSYSYGAIDIK
EPRPEFHAECILOKGELAESEGLFLAOELIADTKERKELSTRVSMLEAIKREK
EHECVNDP"
BASE COUNT      688 a      423 c      461 g      528 t
ORIGIN
Query Match      59.4%; Score 930; DB 1; Length 2100;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1186 TATACAGATGAGAGATTTTAAAGCCAGCAGAGTACAAAATTCGAGAAAATGCCT 1245
DB 1081 TATACAGATGAGAGATTTTAAAGCCAGCAGAGTACAAAATTCGAGAAAATGCCT 1140
OY 1246 CAACCCACATTCAGTGGATGGAGCGAGAAAATAATAGTCTCGATAAAGACCTTCTT 1305
DB 1141 CAACCCACATTCAGTGGATGGAGCGAGAAAATAATAGTCTCGATAAAGACCTTCTT 1200
OY 1306 GGAAGGAGAAATTAAGAAACCCGGCGCTGGGTAATCTGAAAACTCATCTTATAT 1365
DB 1201 GGAAGGAGAAATTAAGAAACCCGGCGCTGGGTAATCTGAAAACTCATCTTATAT 1260
OY 1366 AAGATTAATTAATGATTAATCTCATTGGCCACACCTGCTGGATTAAGCCAGCGCTC 1425
DB 1261 AAGATTAATTAATGATTAATCTCATTGGCCACACCTGCTGGATTAAGCCAGCGCTC 1320
OY 1426 AAGCAGCTGGTATGACCAAAAAACAACCTGCTGATTAATTAATTAATTAATTA 1485
DB 1321 AAGCAGCTGGTATGACCAAAAAACAACCTGCTGATTAATTAATTAATTAATTAATTA 1380
OY 1486 GCTATTGAAGCACTGACCGCTTCATTCAGAAATATGATTCAGTATGACAGCTGCTTA 1545
DB 1381 GCTATTGAAGCACTGACCGCTTCATTCAGAAATATGATTCAGTATGACAGCTGCTTA 1440
OY 1546 GATGACAGCTGCTGATTAATGA 1566
DB 1441 GATGACAGCTGCTGATTAATGA 1461

RESULT 2
LOCUS YPCD1/c 70305 bp DNA BCT 22-MAR-2000
DEFINITION Yersinia pestis plasmid pCD1.
ACCESSION AL117189.1 GI:5832423
VERSION AL117189.1 GI:5832423
KEYWORDS chaperone; cytotoxic effector; IS100; IS1616; IS1617; lcr;
low-calcium response; ysc; targeted effector; transposase; type III
secretion; V antigen; virulence; ylp; yop; ysc.
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis; gamma subdivision; Enterobacteriaceae;
Bacteria.
REFERENCE 1 (bases 1 to 70305)
AUTHORS Karlyshev, A.V. and Wren, B.W.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 70305)
AUTHORS Baker, S.G. and Mungall, K.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 70305)
AUTHORS James, K.D., Parkhill, J., Barrell, B.G. and Randalream, M.A.
JOURNAL Direct Submission
TITLE Submitted (03-SEP-1999) Yersinia pestis sequencing project, Sanger
Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA
COMMENT (URL: http://www.sanger.ac.uk/Projects/Y_pestis/)
Details of Y. pestis sequencing at the Sanger Centre are available
on the World Wide Web.
(UR: http://www.sanger.ac.uk/Projects/Y_pestis/)
CDS are numbered using the following system eg YPCP1.0lc. YP (Y.
pestis), PCP1 (plasmid name), .01 (first CDS), c (complementary
strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given

```


where these have been used to deduce the initiation codon.
 CAUTION: We may not have predicted the correct initiation codon.
 Where possible we choose an initiation codon (atg, gtg, ttg or
 (att)) which is preceded by an upstream ribosome binding site
 sequence (optimally 5-13bp before the initiation codon). If this
 cannot be identified we choose the most upstream initiation codon.

Location/Qualifiers

FEATURES

Source

1. 70305
 /organism="Yersinia pestis"

/plasmid="pCD1"
 /strain="CO-92 Biovar Orientalis"

/db_xref="taxon:632"

repeat_unit

1. 1956
 /note="IS100 element"

gene

88. 1110
 /gene="YPCD1.01"

CDS

88. 1110
 /gene="YPCD1.01"
 /note="YPCD1.01, probable transposase, len: 340 aa;
 putative insertion sequence IS100, identical to
 corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946,
 AF074612) (340 aa), fasta scores: opt: 2328 z-score:
 3808.9 E(): 0, 100.0% identity in 340 aa overlap. Similar
 to many others e.g. TRA0_ECOLI (EMBL:X14793), ISca, E.coli
 transposase for insertion sequence element IS21 (390 aa)
 (33.1% identity in 329 aa overlap). Contains Pfam match to
 entry PF00239 recombinase, site-specific recombinases,
 score 25.70, E-value 4.8e-06. Contains probable
 helix-turn-helix motif at aa 19-40 (Score 2045, +6.15 SD)"

/codon_start=-1
 /transl_table=-1
 /label="YPCD1.01"
 /product="putative transposase"
 /protein_id="CAB54878.1"
 /db_xref="GI:5832424"

/translation="MTWTFVMEKIKIHKQSGSRATARELGSRNTVRYLQAKSP
 PKTPPRAVASLDEYRDYTRQRIADHPKIPRTVARETRDQGRGNTIATFR
 SLSVPDEPAVAFETEPGRQOVWQTMNGSPHFAVLGYSMLIETDNKR
 YDLFETCRNARAFREFGVPREVLYDNNAKTVLQADAYGQHRFHPSLQFGEKMS
 PRLCPREPAQTKGVERVOYTRNSFYIPLMTFLRPMTIVDETANRHLMLHVA
 NORKHETIQARPCDRLMEQSMALPEPEKEDVHDENLVNDKRLPHLPSTIDS
 FCRGVA"

misc_feature

112. 195

/gene="YPCD1.01"

/note="Pfam match to entry PF00239 recombinase,
 site-specific recombinases, score 25.70, E-value 4.8e-06"

gene

1110. 1889

/gene="YPCD1.02"

CDS

1110. 1889

/gene="YPCD1.02"

/note="YPCD1.02, probable transposase, len: 259 aa;
 putative insertion sequence IS100, identical to
 corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946,
 AF074612) (259 aa), fasta scores: opt: 1658 z-score:
 2693.4 E(): 0, 100.0% identity in 259 aa overlap. Similar
 to many others e.g. ISTB_ECOLI (EMBL:X14793), ISB, E.coli
 transposase for insertion sequence element IS21 (265 aa)
 (47.4% identity in 249 aa overlap). Contains P500017
 ATP/GTP-binding site motif A (P-loop)"

/codon_start=-1
 /transl_table=-1
 /label="YPCD1.02"
 /product="putative ATP-binding protein"
 /protein_id="CAB54879.1"
 /db_xref="GI:5832425"

/translation="MMELQHQRLMALAGQLSLISAAPALSOAVDQWSTMDLFE
 HILHEKRLARQKQAMTRAPAKTIEEDFTATGAPKQQLSLSSTLEN
 ENIVLLGPSVGVKTHLAIAAGYAVRAGIVRTTADLLLOLSTAGROORRYTTOR
 GNVAFRLIIDEIGYILFSEQEAKEFQVAVAKRYEKASAMLTSLNPGQDQTFAGDA
 ALTSAMDRLILSHSHVQIKESYRLRKQKAGVIAVANE"

misc_feature

1434. 1457

/gene="YPCD1.02"

/note="P500017 ATP/GTP-binding site motif A (P-loop)"

misc_feature

1956. 46271

/note="identical to Y.pestis KIM5 plasmid pCD1
 (EMBL:AF053946) from 46489 to 1955, and to Y.pestis KIM5
 plasmid pCD1 (EMBL:AF074612) from 59097 to 14563, except
 where noted"

gene

complement(1956..2204)

CDS

complement(1956..2204)

/partial

/gene="YPCD1.03c"

/note="YPCD1.03c, probable transposase remnant, len: 83
 aa, similar to several e.g. N-terminus of TRA0_ECOLI
 (EMBL:X14793), ISca, E.coli transposase for insertion
 sequence element IS21 (390 aa), similarity is interrupted
 by the adjacent IS100 element. The remainder of this CDS
 is in YPCD1.97c. This region is also similar to TR:068707
 (EMBL:AF053946, AF074612) Y.pestis KIM5 putative
 transposase in pCD1 (390 aa). Contains probable
 helix-turn-helix motif at aa 19-40 (Score 1833, +5.43 SD)"

/codon_start=-1

/transl_table=-1

/label="YPCD1.03c"

/protein_id="CAB54880.1"

/db_xref="GI:5832426"

/translation="MISREDFYIKQMRQOQAYIIDATQVGCSEPTVRRYLYKPEPP
 ARKTRKMYLKFMDYIDMRILAENWVNSEVILAEIKAM"

2304. 2948

/gene="YPCD1.04"

2304. 2948

/partial

/gene="YPCD1.04"

/note="YPCD1.04, possible transposase remnant, len: 215
 aa; similar to many e.g. TR:046612 (EMBL:X78052)
 Enterobacter agglomerans IS 1222 ORF8 (276 aa). Truncated
 at N-terminus"

/codon_start=-1

/transl_table=-1

/label="YPCD1.04"

/protein_id="CAB54881.1"

/db_xref="GI:5832427"

/translation="HREGIYHNKRYRYRLYHLSGLVKKRRRRKGLATERPLRPAA
 PNLTRMDVMDLANGRKLCITGVDTKELCYLTVAFGISGVAVTILDSIAFR
 GYPATIRTDQGEPEFTCRALDQWAEVGLRILOPKPONGEIESFNRFDEICNE
 HMFSDVSHARKTISEMRDYNEDRPHSTINTQYPSFEAAMAKGNSDSGSDITK"
 2304. 2982

/note="similar to E. agglomerans IS 1222 (EMBL:X78052) at
 DNA level"

complement(3014..3406)

/gene="YPCD1.05c"

complement(3014..3406)

/gene="YPCD1.05c"

/note="YPCD1.05c, yscE, yera, yope targeting protein, len:
 130 aa; identical to corresponding CDS from Y.pestis KIM5
 pCD1 (EMBL:AF053946, AF074612) (130 aa), fasta scores:
 opt: 861 z-score: 1608.5 E(): 0, 100.0% identity in 130 aa
 overlap and to YERA_YERPE (EMBL:M34279) from Y.pestis
 plasmid pY019. Highly similar to TR:056910 (EMBL:Z18539),
 yscE, from Yersinia enterocolitica (130 aa) (99.2%
 identity in 130 aa overlap) and YERA_YEREN (EMBL:M34278),
 yere, also from Y. enterocolitica (130 aa) (98.5% identity
 in 130 aa overlap). Similar to TR:051448 (EMBL:L27629),
 ORF1, pseudomonas aeruginosa exoenzyme S ORF1 (116 aa)
 (44.3% identity in 115 aa overlap)"

/codon_start=-1

/transl_table=-1

/label="YPCD1.05c"

/product="putative yope chaperone"

/protein_id="CAB54882.1"

/db_xref="GI:5832428"

/translation="MRSFOATITQLRQOLSLSPDIEPIVIGKVGEFACHTHEPVG
 QIMFTPLSDNNDEKETLSNRIFSQDILKPLISDVEGVGAPVLMNRPPLNSLDNS
 LTIQMLTVQGARLQTSLSISPPRSFS"
 3600. 4259

gene

Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 586 ATGATAGAGCTTACGACAAACCAACACATTTTATTTGAGATCTAGAAAAAGTTAGG 645
 |||||
 Db 25311 ATGATTAGAGCTTACGACAAACCAACACATTTTATTTGAGATCTAGAAAAAGTTAGG 25370
 |||||

QY 646 GTGGAACAACCTTACTGTCATGCTTCTTCACGTTTATAGAAAGATTGGTTCACTAGTCAAA 705
 |||||
 Db 25371 GTGGAACAACCTTACTGTCATGCTTCTTCACGTTTATAGAAAGATTGGTTCACTAGTCAAA 25430
 |||||

QY 706 GATTAATAATATAGATTTTCCATTAATATGATCCAGAAAAGATTGGAGTTTGGCC 765
 |||||
 Db 25431 GATTAATAATATAGATTTTCCATTAATATGATCCAGAAAAGATTGGAGTTTGGCC 25490
 |||||

QY 766 AATAGATAATTTACTGATGATATGCAATTTGCTCAAGAAAATCTAGCTTATTTTCTAACCC 825
 |||||
 Db 25491 AATAGATAATTTACTGATGATATGCAATTTGCTCAAGAAAATCTAGCTTATTTTCTAACCC 25550
 |||||

QY 826 GAGGATACCATTTCTTAAAGCGGTCATTTATGACAAACCAACTGCAAAATGCGATCAAGCGA 885
 |||||
 Db 25551 GAGGATACCATTTCTTAAAGCGGTCATTTATGACAAACCAACTGCAAAATGCGATCAAGCGA 25610
 |||||

QY 886 GTAAAGAGTTCCTTGAATATATGCCCAATACAAATGGGAATTCGGGGCTTCATGCGCA 945
 |||||
 Db 25611 GTAAAGAGTTCCTTGAATATATGCCCAATACAAATGGGAATTCGGGGCTTCATGCGCA 25670
 |||||

QY 946 GTATGCAATTTCTTTTAAACCGCGATGTCATGATGATATTTTAAAGTATGTT 1005
 |||||
 Db 25671 GTATGCAATTTCTTTTAAACCGCGATGTCATGATGATATTTTAAAGTATGTT 25730
 |||||

QY 1006 GATTCATTAATGATTCATGATGATGATGCCGATGACAGATTCGCTGAAGAATTAAGTACGAGCTT 1065
 |||||
 Db 25731 GATTCATTAATGATTCATGATGATGATGCCGATGACAGATTCGCTGAAGAATTAAGTACGAGCTT 25790
 |||||

QY 1066 ACCGCCAATTAAGATTTATTCAGTTATTTAAGCCGAATTAATTAAGCATCTGTCAGT 1125
 |||||
 Db 25791 ACCGCCAATTAAGATTTATTCAGTTATTTAAGCCGAATTAATTAAGCATCTGTCAGT 25850
 |||||

QY 1126 AGTGGACACATTAATTCATGATGATTAATTCATTTATTCATGATGATTAATTTATATGCT 1185
 |||||
 Db 25851 AGTGGACACATTAATTCATGATGATTAATTCATTTATTCATGATGATTAATTTATATGCT 25910
 |||||

QY 1186 TATACAGATGAAGATTTTAAAGCGCAGAGATGACAAATTCCTGAGAAAATGCGCT 1245
 |||||
 Db 25911 TATACAGATGAAGATTTTAAAGCGCAGAGATGACAAATTCCTGAGAAAATGCGCT 25970
 |||||

QY 1246 CAATCCACCATTCAGTGGATGAGGACGAGAAAATATAGTCTGATTAAGGACTTTCTT 1305
 |||||
 Db 25971 CAATCCACCATTCAGTGGATGAGGACGAGAAAATATAGTCTGATTAAGGACTTTCTT 26030
 |||||

QY 1306 GGAAGTGAATTAAGAAACCGGGCGTTGGTATATCGAAAACTCATCTATATAT 1365
 |||||
 Db 26031 GGAAGTGAATTAAGAAACCGGGCGTTGGTATATCGAAAACTCATCTATATAT 26090
 |||||

QY 1366 AAGATATAATGATTAATTCATCTTGGCCACCACTCTCGGATAAGTCCAGGCGCTC 1425
 |||||
 Db 26091 AAGATATAATGATTAATTCATCTTGGCCACCACTCTCGGATAAGTCCAGGCGCTC 26150
 |||||

QY 1426 AAGCATTTGGTTAGCCAAAAAACAACACAGCTGCTGATTAATCATCAGTTTAATCA 1485
 |||||
 Db 26151 AAGCATTTGGTTAGCCAAAAAACAACACAGCTGCTGATTAATCATCAGTTTAATCA 26210
 |||||

QY 1486 GCTTTTGAAGACTGACACCGCTTTCATTCAGAAATATGATTAAGTATGCAACGCTGCTA 1545
 |||||
 Db 26211 GCTTTTGAAGACTGACACCGCTTTCATTCAGAAATATGATTAAGTATGCAACGCTGCTA 26270
 |||||

QY 1546 GATGACACGCTGCTGATAATGA 1566
 |||||
 Db 26271 GATGACACGCTGCTGATAATGA 26291
 |||||

RESULT 4
 AF074612

LOCUS AF074612 70559 bp DNA circular BCT 07-APR-2000
 DEFINITION Yersinia pestis plasmid pCD1, complete plasmid sequence.
 ACCESSION AF074612 M25810
 VERSION AF074612.1 GI:3822037
 KEYWORDS Yersinia pestis.
 ORGANISM Yersinia pestis
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.

REFERENCE
 AUTHORS Leung,K.Y. and Straley,S.C.
 TITLE The yopM gene of Yersinia pestis encodes a released protein having homology with the human platelet surface protein GPIb alpha J. Bacteriol. 171 (9), 4623-4632 (1989)
 JOURNAL 89359090
 MEDLINE
 REFERENCE 2 (bases 1 to 70559)
 AUTHORS Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and Blattner,F.R.
 TITLE DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia pestis KIM5 Infect. Immun. 66 (10), 4611-4623 (1998)
 JOURNAL 98427122
 MEDLINE
 REFERENCE 3 (bases 43318 to 44840)
 AUTHORS Straley,S.C.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-1993) Microbiology and Immunology, University of Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA
 REFERENCE 4 (bases 1 to 70559)
 AUTHORS Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and Blattner,F.R.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUN-1998) Microbiology and Immunology, University of Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA
 COMMENT On Apr 7, 2000 this sequence version replaced gi:1155539.
 FEATURES
 source
 1. 70559
 /organism="Yersinia pestis"
 /plasmid="pCD1"
 /strain="KIM5"
 /db_xref="taxon:632"
 57. 368
 /gene="Y0001"
 57. 368
 /gene="Y0001"
 /note="O103: 43 pct identical (0 gaps) to 100 residues of an approx. 200 aa protein GENEPT: g1:537126, ori_0198 Escherichia coli"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAC69758.1"
 /db_xref="GI:3822038"
 /translation="MHQOSKAGASRTLSILMRQSGYNNVRLARLRECGLASROP GKPRYGEREVSILASPDLLKROKRPSEPNRVMSGYISYIKVNGWCYIALVIDLYSH W"
 665. 1033
 /gene="nuc"
 /note="Y0002"
 665. 1033
 /gene="nuc"
 /codon_start=1
 /transl_table=11
 /product="endonuclease"
 /protein_id="AAC69759.1"
 /db_xref="GI:3822039"
 /translation="MDTKLLOHTPIGTMVDYRPVNTKSGKRLRRCDFVIHYRMDL VNAGIVRVNSFKALHDKVILYIDGANTQMGSRNFSQAAVQNSENVLLIMGFTVVO AIIQYQSRNKKGTWRSSTY"
 1171. 1425
 /gene="repB"
 /note="Y0003"
 1171. 1425
 /gene="repB"

```

misc_RNA
    /codon_start-1
    /transl_table-11
    /product-"replication protein"
    /protein_id-"AAC69760.1"
    /db_xref="GI:3822040"
    /translation-"MSQIDNATSSSKRAYRKGNPLTGAEKQMSVSRKKEHKAIV
    FIDNDLKNELDLEDGSGLTQTEMERWQREKAATNA"
    complement(1560..1649)
    /note-"antisense RNA"
    /product-"copa"
    1667..1741
    /gene-"tap"
    /note-"y0004"
    1667..1741
    /gene-"tap"
    /codon_start-1
    /transl_table-11
    /product-"repa translation protein"
    /protein_id-"AAC69761.1"
    /db_xref="GI:3822041"
    /translation-"MFRKQVYLLRLLLPCNISAGRCD"
    1734..2600
    /gene-"repa"
    /note-"y0005"
    1734..2600
    /gene-"repa"
    /codon_start-1
    /transl_table-11
    /product-"replication protein"
    /protein_id-"AAC69762.1"
    /db_xref="GI:3822042"
    /translation-"MNHQALFTHHYQYKNPNPEPTPEGKRTLPFCCKLMAKEG
    TREFSMHVAFAARSILSRHMPPLLRRAIDALOGMCFHDPPLANRQSRITNAI
    ECGLATESKGNMSTIRATRALRFLSELGLTYOTEDYDQICNIPDTITTPALFSA
    IDVSDVAVAARBSRVRMENQCRKORLRLMDLAIAMRFYRERPSYTERKAH
    GIKRARADVDTRRDITAIYNRQITREIAEGRFGNIDAVREKARAKVKKEMLSR
    NNNYRLATGAT"
    1741..2935
    /note-"oriR"
    complement(3427..3645)
    /gene-"y0006"
    complement(3427..3645)
    /gene-"y0006"
    /note-"OriF (f72): 42 pct identical (0 gaps) to 33
    residues of an approx. 216 aa protein GENEPT: g112055297,
    Imp2. Xenopus laevis"
    /codon_start-1
    /transl_table-11
    /product-"unknown"
    /protein_id-"AAC69763.1"
    /db_xref="GI:3822043"
    /translation-"MRSPVAGCSYTVVTVQLNHDHKSAYKHKRLYHVLGQASS
    AQPVSMPHKGKAPADINYSVDNKR"
    4758..5186
    /gene-"y0008"
    4758..5186
    /gene-"y0008"
    /note-"OriF (ol42): 31 pct identical (1 gap) to 48
    residues of an approx. 104 aa protein GENEPT: g112149940,
    Ori1. P. syringae"
    /codon_start-1
    /transl_table-11
    /product-"unknown"
    /protein_id-"AAC69764.1"
    /db_xref="GI:3822044"
    /translation-"MINTFTLPRKASHFGIDRLSODEYGLCELLINDRVIMLRA
    DELNKLTLGLPILGFSGEFAASAOLFECYSINALNKGCFAMSEELGILAEKHL
    SLDELNENVSSEIANFYDMLSLVSLPATAFSYSISIG"
    5204..7402
    /gene-"ypka"
    /note-"y0009"
    5204..7402
    /gene-"ypka"

```

```

gene
    /codon_start-1
    /transl_table-11
    /product-"secreted protein kinase"
    /protein_id-"AAC69765.1"
    /db_xref="GI:3822045"
    /translation-"MKSVMKIMTPPSSISLAKAHERISQHWNPYGEINIGKRYRI
    DNQVILNPHSGFSLFREGVGKIFSGKFNFSIANLIDTLAAAKTISOELRSDIPN
    ALSNLFQAKPQTELPGLGWKEPLSGAPLEGRVAVETDFAEGESHISIIETKQRL
    VAKIERISIAEGHLPALFAEKHYIYTAGKHPLAVHGMNAVYPGNRKEEALIMDEVD
    GMRGSDTLRLTADSKOKGINSKINSEAYWGTIKFAHRLIDVTNHLAKAGVNDIKRGN
    VEDRASGPVYIDGLHSRSGEQPKGFIESKAPBLGVCNIGASKRSYFLVSTLLH
    CIEGFENKPEIKPNQGLRITSEPAHVDENGYPIHRPGIAGVETAYTRFTDILGVS
    ADSRPSNEARLHEFLSDGTIDESAKQILKDTLGEPSLSTDVARTTPKRLSLD
    LLRHLSSAATKQIDMGVLSLDLMLVALKAEREGGVKQDLKFSNLILKTYRVI
    EDYVKGREGDTKNSSTSEVSPYRSNFMLSIYEPISLORLOKHLDOTHSFSDIGSLVRAH
    KHELELLVIVTLISQGGQPVSEETGYFLNRLTEAKITISQOLNTIQCOOESAKADSI
    LNRSGSNADVARSQSLQRDSTQPVYKGTEDQYTAIRHQMMAHAHATTLQVSEFTDD
    MNFTVDSIPLLIQLGRSSLMDEHLVEQREKIRELTITTAERLRLERBM"
    7798..8664
    /gene-"yopJ"
    /note-"y0010"
    7798..8664
    /gene-"yopJ"
    /codon_start-1
    /transl_table-11
    /product-"targeted effector protein"
    /protein_id-"AAC69766.1"
    /db_xref="GI:3822046"
    /translation-"NIGPISQINISGGISEKETSLSJSEBKNIITOLETIDSGSW
    FHNKSRMDVEVPALVLOANNKYPENMLNVTSLDLSIEIKNVIENGVSRRPIN
    MEGGIIHPSVIDYKHNGKTSILPEPANFNMSGAMARIKRTAIEYNGVSSRPIIN
    VMEDIORSSSECGIFSLAKLKYIERPSLKIHDNDKGLISDENLPHDKIDPYL
    PVTFKHQGGKRLMEYINTNQGQGVYNNKNEITVAFPNKNSIVDGKELSVYKH
    KRIAEYKILLYK"
    9858..10064
    /gene-"y0012"
    9858..10064
    /gene-"y0012"
    /note-"O68: 45 pct identical (0 gaps) to 68 residues of an
    approx. 560 aa protein GENEPT: g112162435, hypothetical"
    /codon_start-1
    /transl_table-11
    /product-"unknown"
    /protein_id-"AAC69767.1"
    /db_xref="GI:3822047"
    /translation-"YAVGRKNLPLGSLRAGORNASILSLETAKLNHDPYWLRYD
    LRLPTWPSQNLALLPYAENRFS"
    10347..11753
    /gene-"yopH"
    /note-"y0013"
    10347..11753
    /gene-"yopH"
    /codon_start-1
    /transl_table-11

```

Query Match 59.4%; Score 930; DB 2; Length 70559;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 586 ANGATTAGAGCCTACGACAAAACCCAAACATTTATAGAGACTAGAAAAAGTTAG 645
    |||
DB 37919 ATGATTAGAGCCGACGACAAAACCCAAACATTTATATGAGAGCTGAAAAAGTTAG 37978
    |||
QY 646 GTGGAAACAACCTACTGCTATGGTCTTCAAGTTTGAAGAATGGTTCAGTTAGTCAAA 705
    |||
DB 37979 GTGGAAACAACCTACTGCTATGGTCTTCAAGTTTGAAGAATGGTTCAGTTAGTCAAA 38038
    |||
QY 706 GATTAATAATATGATTTCCATTAATATGATCCAGAAAAGATTCGAGATTTTGGCC 765
    |||
DB 38039 GATTAATAATATGATTTCCATTAATATGATCCAGAAAAGATTCGAGATTTTGGCC 38098
    |||
QY 766 AATAGAGTAATTACTGATGATTCGATTCGACAGAAAATCCTAGCTATTTCTACCC 825
    |||

```

Db 38099 AATGAGTATTACTGATGATATTCGAAATTCCTCAGAAAATCTAGCTATTCTACCC 38158
 QY 826 GAGATACCATCTCTTAAGCGGTCATTATGACAAACCAATGCAAAATGGCATCAAGCA 885
 Db 38159 GAGGATGCCATTCTTAAGCGGTCATTATGACAAACCAATGCAAAATGGCATCAAGCA 38218
 QY 886 GTAAAAAGATCTCTGATATCATCGCCGAATACAAATGGGAATGGCGGCTCATGGCA 945
 Db 38219 GTAAAAAGATCTCTGATATCATCGCCGAATACAAATGGGAATGGCGGCTCATGGCA 38278
 QY 946 GTAAATGATCTCTCTTAACCGCCGATCTATCGATGATGATTTTGAAGATGTT 1005
 Db 38279 GTAAATGATCTCTCTTAACCGCCGATCTATCGATGATGATTTTGAAGATGTT 38338
 QY 1006 GATTCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
 Db 38339 GATTCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 38398
 QY 1066 ACCGCCGAATTAAGATTTATTCAGTTATTCAGCCGAATTAATAGCATCTGTCTAGT 1125
 Db 38399 ACCGCCGAATTAAGATTTATTCAGTTATTCAGCCGAATTAATAGCATCTGTCTAGT 38458
 QY 1126 AATGGCACCATAATATTCATGATTAATTCATGATGATGATGATGATGATGATGATGAT 1185
 Db 38459 AATGGCACCATAATATTCATGATTAATTCATGATGATGATGATGATGATGATGATGAT 38518
 QY 1186 TATACAGATGAAGAGATTTTAAAGCCAGCGAGATGACAAATTCGAGAAAATGCGT 1245
 Db 38519 TATACAGATGAAGAGATTTTAAAGCCAGCGAGATGACAAATTCGAGAAAATGCGT 38578
 QY 1246 CAACCCACATTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
 Db 38579 CAACCCACATTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 38638
 QY 1306 GGAATGAGAAATTAAGAAACCGGCGGTGATGATGATGATGATGATGATGATGATGATGAT 1365
 Db 38639 GGAATGAGAAATTAAGAAACCGGCGGTGATGATGATGATGATGATGATGATGATGATGAT 38698
 QY 1366 AAGATTAATTAATGATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425
 Db 38699 AAGATTAATTAATGATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 38758
 QY 1426 AAGCATGCTGTAGCCAAAAAACAATCACTGCTGTGATTAATACATCAAGTTTAAATCA 1485
 Db 38759 AAGCATGCTGTAGCCAAAAAACAATCACTGCTGTGATTAATACATCAAGTTTAAATCA 38818
 QY 1486 GCTATTGAAGCACTGACCGCTTCAATCAAGAAATATGATGATGATGATGATGATGATGATGAT 1545
 Db 38819 GCTATTGAAGCACTGACCGCTTCAATCAAGAAATATGATGATGATGATGATGATGATGATGAT 38878
 QY 1546 GATGACAGCTGTGTAATGA 1566
 Db 38879 GATGACAGCTGTGTAATGA 38899

RESULT 5
 A46411
 LOCUS A46411 1014 bp DNA PAT 07-MAR-1997
 DEFINITION Sequence 1 from Patent WO9524475.
 ACCESSION A46411
 VERSION A46411.1 GI:2300612
 KEYWORDS
 SOURCE
 ORGANISM
 Yersinia pestis.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.
 REFERENCE
 1 (bases 1 to 1014)
 AUTHORS Titball, R.W., Williamson, E.D. and Leary, S.E.
 TITLE VACCINE COMPOSITIONS
 JOURNAL Patent: WO 9524475-A 1 14-SEP-1995;
 COMMENT
 FEATURES
 Other publication AU 1853995 950925.
 Location/Qualifiers

source 1..1014
 /organism="Yersinia pestis"
 /db_xref="taxon:632"
 CDS 1..990
 /note="unamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA02893.1"
 /db_xref="GI:2300613"
 /translation="MSERIRAEONPQHFIEDLEKRVLEQTLGHSSTLEELVQLVKD
 KNIDISITDKRSEVRAAVITDIDLELKLILATFPEPAILLGGHYDNOLOGIK
 RVEFLESSPTQWELRAFMVMEHLADRIDDIILVIVDSNMHSDASKLEEL
 AELTAELKIVSVIQAELINKHLSGGTINIHDKSLMDKNLYGYDEEIPFASAEYKI
 LEKPOFTIOYDSEKRVSLKIDFSGSNKRGALGNKNSYKKNELSHFATTC
 SDKSRPLNDLVISOKTTOUSDITSRNSAIEALNRIQKVDVMDRLDDTSK"
 BASE COUNT 346 a 181 c 201 g 286 t
 ORIGIN
 Query Match 59.2%; Score 927; DB 5; Length 1014;
 Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;
 Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 589 ATTAGAGCTAGAACAAAACCAACATTTTATGAGATCTAGAAAAGTTAGGCTG 648
 Db 13 ATTAGAGCTAGAACAAAACCAACATTTTATGAGATCTAGAAAAGTTAGGCTG 72
 QY 649 GAACAACCTACGTCATGCTGCTTCAGTTTGAAGAATTGGTCAGTACCAAGAT 708
 Db 73 GAACAACCTACGTCATGCTGCTTCAGTTTGAAGAATTGGTCAGTACCAAGAT 132
 QY 709 AAAAATATGATATTTCCATTAATATGATGCCAGAAAAGTTGGAGGTTTTGCCAAT 768
 Db 133 AAAAATATGATATTTCCATTAATATGATGCCAGAAAAGTTGGAGGTTTTGCCAAT 192
 QY 769 AGAGTAATTAATGATGATATTCGAATTCCTCAAGAAAATCTAGCTATTTTACCCGAG 828
 Db 193 AGAGTAATTAATGATGATATTCGAATTCCTCAAGAAAATCTAGCTATTTTACCCGAG 252
 QY 829 GATACCATCTTAAAGCGGCTATTATGACAACAATGCAAAATGGCATCAAGCAGTA 888
 Db 253 GATACCATCTTAAAGCGGCTATTATGACAACAATGCAAAATGGCATCAAGCAGTA 312
 QY 889 AAGAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 Db 313 AAGAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
 QY 949 ATGCATTTCTTTAAACCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
 Db 373 ATGCATTTCTTTAAACCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
 QY 1009 TCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
 Db 433 TCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
 QY 1069 GCCGAAATTAAGATTTATCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
 Db 493 GCCGAAATTAAGATTTATCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 552
 QY 1129 GCCGAAATTAAGATTTATCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
 Db 553 GCCGAAATTAAGATTTATCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 612
 QY 1189 ACAGATGAAGAGATTTTAAAGCCAGCGAGATGATGATGATGATGATGATGATGATGATGAT 1248
 Db 613 ACAGATGAAGAGATTTTAAAGCCAGCGAGATGATGATGATGATGATGATGATGATGATGAT 672
 QY 1249 ACCACATTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
 Db 673 ACCACATTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
 QY 1309 ACTGAGAAATTAAGAACCGGCGCTGGTGAATCTGAAAATCTACTCTATATATAA 1368

```

Db 733 AGTGAATTAAGAAACACCGGGCGTTGGTATCTGAAAACTCATCTTTAATAA 792
OY 1369 GATAATTAATGATTAATCTCACTTTCACACACCTGCTCGATAAGTCAGGCCCTCAAC 1428
Db 793 GATTAATTAATGATTAATCTCACTTTCACACACCTGCTCGATAAGTCAGGCCCTCAAC 852
OY 1429 GACTGTGTTGCCCAAAAACACACCTGCTGCTATATATATATATATATATATATAT 1488
Db 853 GACTGTGTTGCCCAAAAACACACCTGCTGCTATATATATATATATATATATATATAT 912
OY 1489 ATTGAACACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGCTAGAT 1548
Db 913 ATTGAACACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGCTAGAT 972
OY 1549 GACACGCTCTGTAATGA 1566
Db 973 GACACGCTCTGTAATGA 990

RESULT 6
LOCUS A56793 1014 bp DNA PAT 03-MAR-1998
DEFINITION Sequence 1 from Patent WO9628551.
ACCESSION A56793
VERSION A56793.1 GI:3712808
KEYWORDS
SOURCE
ORGANISM Yersinia pestis.
Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Tibbitts, R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C.,
Bennett and Alice, M.
TITLE VACCINES FOR PLAGUE
JOURNAL Patent: WO 9628551-A 1 19-SEP-1996;
COMMENT Other publication ZA 9602036 960716
Other publication AU 4951196 961002.
FEATURES
source
1. 1014
/organism="Yersinia pestis"
/db_xref="taxon:632"
<1.990
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA03419.1"
/db_xref="GI:4530021"
/translation="ISEFTIRAYEQNHFIEDLEKRVEDLTGHSSVLEELYQVLYKD
KNDISIKYDPRKDSVFANRVITDDIELKTLAYFLEPDALDKGHDNOMONIK
RKVEFLSPNTOVRAFMVAHNSFLTRIDDDILKLYVSMNHGHRASRLREL
AELTALIKIYVIOAEIKNKHSSTGTLINHDKISINMDKLYGTDEIEFSAEYKI
LERMPOTIOVDSEKKIYISINDFISGNRRGALINLKNYSYNNDNNEISFATTC
SPKSRPLNDLVSKTQLOSDIYSRENSAIFALNRFQIKIDSVQRLLDTSGR"
CDS
<1.990
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA03419.1"
/db_xref="GI:4530021"
/translation="ISEFTIRAYEQNHFIEDLEKRVEDLTGHSSVLEELYQVLYKD
KNDISIKYDPRKDSVFANRVITDDIELKTLAYFLEPDALDKGHDNOMONIK
RKVEFLSPNTOVRAFMVAHNSFLTRIDDDILKLYVSMNHGHRASRLREL
AELTALIKIYVIOAEIKNKHSSTGTLINHDKISINMDKLYGTDEIEFSAEYKI
LERMPOTIOVDSEKKIYISINDFISGNRRGALINLKNYSYNNDNNEISFATTC
SPKSRPLNDLVSKTQLOSDIYSRENSAIFALNRFQIKIDSVQRLLDTSGR"
BASE COUNT 346 a 181 c 201 g 286 t
ORIGIN
Query Match 59.2%; Score 927; DB 5; Length 1014;
Best local Similarity 99.9%; Pred. No. 0;
Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 769 AGATAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 828
Db 193 AGATAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 252
OY 829 GATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 888
Db 253 GATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 312
OY 889 AAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
Db 313 AAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
OY 949 ATGATTTCTCTTTAACCGCCGATGATGATGATGATGATGATGATGATGATGATGAT 1008
Db 373 ATGATTTCTCTTTAACCGCCGATGATGATGATGATGATGATGATGATGATGATGAT 432
OY 1009 TCAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
Db 433 TCAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
OY 1069 GCGGAATTAAGATTTATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 1128
Db 493 GCGGAATTAAGATTTATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 552
OY 1129 GCGACATTAATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 1188
Db 553 GCGACATTAATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 612
OY 1189 ACAGATTAAGATTTTAAAGCCAGGAGTACAAATTTCTCGAAGAAATGCTCAA 1248
Db 613 ACAGATTAAGATTTTAAAGCCAGGAGTACAAATTTCTCGAAGAAATGCTCAA 672
OY 1249 ACCGACATTAAGATTTTAAAGCCAGGAGTACAAATTTCTCGAAGAAATGCTCAA 1308
Db 673 ACCGACATTAAGATTTTAAAGCCAGGAGTACAAATTTCTCGAAGAAATGCTCAA 732
OY 1309 AGTGAATTAAGAAACCGGGCGTTGGTATCTGAAAACTCATACTTTAATAA 1368
Db 733 AGTGAATTAAGAAACCGGGCGTTGGTATCTGAAAACTCATACTTTAATAA 792
OY 1369 GATAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1428
Db 793 GATAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 852
OY 1429 GACTGTGTTGCCCAAAAACACACCTGCTGCTATATATATATATATATATATATAT 1488
Db 853 GACTGTGTTGCCCAAAAACACACCTGCTGCTATATATATATATATATATATATATAT 912
OY 1489 ATTGAACACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGCTAGAT 1548
Db 913 ATTGAACACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGCTAGAT 972
OY 1549 GACACGCTCTGTAATGA 1566
Db 973 GACACGCTCTGTAATGA 990

RESULT 7
LOCUS A56808 1462 bp DNA PAT 03-MAR-1998
DEFINITION Sequence 16 from Patent WO9628551.
ACCESSION A56808
VERSION A56808.1 GI:3712821
KEYWORDS
SOURCE
ORGANISM Yersinia pestis.
Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 1462)
AUTHORS Tibbitts, R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C.,
Bennett and Alice, M.
TITLE VACCINES FOR PLAGUE

```

JOURNAL

Patent: WO 9628551-A 16 19-SEP-1996;

COMMENT

SECR DEFENCE (GB)
Other publication ZA 9602036 960716

Other publication AU 4951196 961002.

FEATURES

Location/Qualifiers

1. 1462

/organism="Yersinia pestis"

/db_xref="taxon:632"

<8. 1450

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CA03421.1"

/db_xref="GI:3712822"

/translation="ADLFASSTTATNTLVEPARITLYREGAPITINDGNIDELVLC

TLVIGKRTGTSVNFETDAGDPMTLFTSODGNHOFETKVGKDSRDPDISPKV

NGENLVGDVVLATGSDPEFVRSIGSKGLAAGKATVAVTVSNOSIGRLRAYE

ONPQHFIEDLEKRVLEQSGSVLEELVOLKRNKIDISIKYPRDSVFARVY

IDDIELKRLIAYFLPEDALILKGHYDQNLQIKRKEPLESSSPYQWELRAFAVM

HESLADRIDDDILKVIYDSNMHNGDASKIREELAEETAEKIKYSYQAEIKHLS

SGTINHRKSNINMDKNTXGTDEIEFRASAEKYLEMPOTTIOVDSSEKIVSIKD

FLGSENRKRTGALGNLKNKSYINKDNELSHFATTCSDKSRPLNDLVSKTQULSDITS

RNSAIEALNFRIOKRYDSVMORLDDTISGK"

BASE COUNT 476 a 286 c 300 g 400 t

ORIGIN

Query Match 59.2%; Score 927; DB 5; Length 1462;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 589 ATTAGAGCTAGAGCAAAACCCACACATTTATTGAGAGCTAGAAAAGTAGGGG 648
 Db 473 ATTAGAGCTAGAGCAAAACCCACACATTTATTGAGAGCTAGAAAAGTAGGGG 532
 Oy 649 GAACAACCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 708
 Db 533 GAACAACCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 592
 Oy 709 AAAAAATATAGATATTTCCATTAATATGATCCAGAAAAGTAGGGGTTTGGCAAT 768
 Db 593 AAAAAATATAGATATTTCCATTAATATGATCCAGAAAAGTAGGGGTTTGGCAAT 652
 Oy 769 AGAGATATCTAGATGATTCGAATTCGTCAGAAAAGTAGGGGTTTGGCAAT 828
 Db 653 AGAGATATCTAGATGATTCGAATTCGTCAGAAAAGTAGGGGTTTGGCAAT 712
 Oy 829 GATACCATCTTAAAGCGCGGTATGACAAACCACTGCAAAATGGCATCAAGGAGTA 888
 Db 713 GATACCATCTTAAAGCGCGGTATGACAAACCACTGCAAAATGGCATCAAGGAGTA 772
 Oy 889 AAGAGATCTTGAATCATGCGCGAATACACATGGAATGCGGCGTTGATGCGAGTA 948
 Db 773 AAGAGATCTTGAATCATGCGCGAATACACATGGAATGCGGCGTTGATGCGAGTA 832
 Oy 949 ATGCATTTCTTTAAACCCCGCATGATGATGATGATGATGATGATGATGATGATG 1008
 Db 833 ATGCATTTCTTTAAACCCCGCATGATGATGATGATGATGATGATGATGATGATG 892
 Oy 1009 TCAATGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1068
 Db 893 TCAATGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952
 Oy 1069 GCCGATTAAGATTTATCACTTATTCAGCGGAATTAATAGCATGTCTAGTAGT 1128
 Db 953 GCCGATTAAGATTTATCACTTATTCAGCGGAATTAATAGCATGTCTAGTAGT 1012
 Oy 1129 GGCACATTAATATCCATATTAATCCATTAATTCATGATTAATTAATGATGAT 1188
 Db 1013 GGCACATTAATATCCATATTAATCCATTAATTCATGATTAATTAATGATGAT 1072
 Oy 1189 ACAGATGAAGATTTTAAACCGCGAGATGACAAAATCTGAGAAATGCTCTCA 1248

Db 1073 ACAGATGAAGAGATTTTAAACCGCGAGATGACAAAATCTGAGAAATGCTCTCA 1132
 Oy 1249 ACCACATTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1308
 Db 1133 ACCACATTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1192
 Oy 1309 AGTGAATTAAGAAACCGGCGGTGATGATGATGATGATGATGATGATGATGATG 1368
 Db 1193 AGTGAATTAAGAAACCGGCGGTGATGATGATGATGATGATGATGATGATGATG 1252
 Oy 1369 GATTAATTAAGATTTATCTACCTTTCACACCTGCTGATGATGATGATGATGATG 1428
 Db 1253 GATTAATTAAGATTTATCTACCTTTCACACCTGCTGATGATGATGATGATGATG 1312
 Oy 1429 GACTGGTTAGCCAAAACCACTGATGATGATGATGATGATGATGATGATGATG 1488
 Db 1313 GACTGGTTAGCCAAAACCACTGATGATGATGATGATGATGATGATGATGATG 1372
 Oy 1489 ATTGAAGCACTGACCGTTTCAATGAAATGATGATGATGATGATGATGATGATG 1548
 Db 1373 ATTGAAGCACTGACCGTTTCAATGAAATGATGATGATGATGATGATGATGATG 1432
 Oy 1549 GACAGCTCTGTAATGA 1566
 Db 1433 GACAGCTCTGTAATGA 1450

RESULT 8 A56814 1530 bp DNA PAT 03-MAR-1998

LOCUS A56814 Sequence 22 from Patent WO9628551.

DEFINITION A56814.1 GI:3712827

ACCESSION A56814.1

VERSION A56814.1

KEYWORDS Yersinia pestis.

SOURCE Yersinia pestis.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Yersinia.

REFERENCE 1 (bases 1 to 1530)

AUTHORS Tildall,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,

Bennett and Alice,M.

TITLE VACCINES FOR PLAGUE

JOURNAL PATENT: WO 9628551-A 22 19-SEP-1996;

COMMENT SECR DEFENCE (GB)

Other publication ZA 9602036 960716

Other publication AU 4951196 961002.

FEATURES Location/Qualifiers

SOURCE 1. 1530

/organism="Yersinia pestis"

/db_xref="taxon:632"

13. 1518

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CA03423.1"

/db_xref="GI:3712828"

/translation="MKKISSVIAIALFETIATANAADLTASTATATLVEPARITLY

KEGAPITMDNGNIDTELVTGLIGKRTGTSVNFETDAGDPMTLFTSODGNHOFETKVGK

DSRDPDISPKVNGENLVGDVVLATGSDPEFVRSIGSKGLAAGKATVAVTVSNOSIGRLRAYE

ONPQHFIEDLEKRVLEQSGSVLEELVOLKRNKIDISIKYPRDSVFARVY

IDDIELKRLIAYFLPEDALILKGHYDQNLQIKRKEPLESSSPYQWELRAFAVM

HESLADRIDDDILKVIYDSNMHNGDASKIREELAEETAEKIKYSYQAEIKHLS

SGTINHRKSNINMDKNTXGTDEIEFRASAEKYLEMPOTTIOVDSSEKIVSIKD

FLGSENRKRTGALGNLKNKSYINKDNELSHFATTCSDKSRPLNDLVSKTQULSDITS

RNSAIEALNFRIOKRYDSVMORLDDTISGK"

BASE COUNT 500 a 296 c 311 g 423 t

ORIGIN

Query Match 59.2%; Score 927; DB 5; Length 1530;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	1133	CCATTAATAATCCAGATTTAAACCATTTATCTCAGTGAATAAATTATATGTTATACAG	1192
Db	557	CCATTAATAATCCAGATTTAAACCATTTATCTCAGTGAATAAATTATATGTTATACAG	616
OY	1193	ATGAAGACATTTTTAAAGCCAGCGCACAGTAGTCMAAATTCCTGAGAAAAATGCCCAACA	1252
Db	617	ATGAAGAATTTTTAAAGCCAGCGCACAGTAGTCMAAATTCCTGAGAAAAATGCCCAACA	676
OY	1253	CCATTTCAGGTGGATGGGAGCGAGAAAAAATATAGTCTGCATTAAGSACTTCTTGGAATG	1312
Db	677	CCATTTCAGGTGGATGGGAGCGAGAAAAAATATAGTCTGCATTAAGSACTTCTTGGAATG	736
OY	1313	AGAAATAAAAGAACCGGGGCGCTGGGTGATCTGAAAAAATCTACTATATATAAGATA	1372
Db	737	AGAAATAAAAGAACCGGGGCGCTGGGTGATCTGAAAAAATCTACTATATATAAGATA	796
OY	1373	ATAATGAATTAATCTCACTTTTCCACACACCCTCTCGGATAGTCCAGGCCGCTCAACGACT	1432
Db	797	ATAATGAATTAATCTCACTTTTCCACACACCCTCTCGGATAGTCCAGGCCGCTCAACGACT	856
OY	1433	TGGTTAGCCAAAAAACAACACTCAGCTGTGTGATATTACATCACGTTTATATTCAGTATG	1492
Db	857	TGGTTAGCCAAAAAACAACACTCAGCTGTGTGATATTACATCACGTTTATATTCAGTATG	916
OY	1493	AAGCAGTAAACCGTTTCATTCAGAAATATGATTCGATATGCAACGCTCTGATAGTACA	1552
Db	917	AAGCAGTAAACCGTTTCATTCAGAAATATGATTCGATATGCAACGCTCTGATAGTACA	976
OY	1553	CGCTGTGTAATGA	1566
Db	977	CGCTGTGTAATGA	990
RESULT	10		
LOCUS	A56795		
DEFINITION	Sequence 3 from Patent WO9628551.	PAT	03-MAR-1998
ACCESSION	A56795		
VERSION	A56795.1	GI:3712810	
KEYWORDS			
SOURCE	Yersinia pestis.		
ORGANISM	Yersinia pestis.		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
AUTHORS	Yersinia.		
TITLE	1 (bases 1 to 1014)		
JOURNAL	Tibball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C., Bennett, and Alice,M.		
COMMENT	VACCINES FOR PLAGUE.		
FEATURES	Patent: WO 9628551-A 3 19-SEP-1996;		
source	SECR DEFENCE (GB) 2A 9602036 960716		
	Other publication AU 4951196 961002.		
CDS	Location/Qualifiers		
	1..1014		
	/organism="Yersinia pestis"		
	/db_xref="taxon:632"		
	<1..990		
	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA03420.1"		
	/db_xref="GI:3712811"		
	/translation="GPEIRAYEDNPQHFIEDLEKRVBEOLITGHSSVLELVQVKD		
	KNIIDISIKPYDRKDSEVFANVTDDIELKLILAYFLPEDAILKGHYNDLNGLIK		
	RVEKFELSPNTOWELRAFMVAHMFSLIADIDDIILKIVDSNNHHGDARSKEBEL		
	ALFELAETKIVIOAEINKHLSSGGINIHDKSNIMDNKNLYGTDEIFKASLEYRI		
	LEKHPOPTIOVDGESEKKIVSIKDFLSENKRRTGALGNKNSYSNKNDELSHPATTC		
	SDKRPLNDLVSKRTQSLLDTLSRFSNAIEALNFIDKIYDSVMRLDDDISGR"		
BASE COUNT	343 a	184 c	205 g 282 t
ORIGIN			

Query Match		58.9%;	Score 923;	Db 5;	Length 1014;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 973;		Conservative	0;	Mismatches 1;	Indels 0;
					Gaps 0;
QY	593	GAGCCTACGAAACAAACCACACACATTTTATGTAGAGATCTAGAAAAAGTTAGGGTGAGAC	652		
Db	17	GAGCCTACGAAACAAACCACACACATTTTATGTAGAGATCTAGAAAAAGTTAGGGTGAGAC	76		
QY	653	AACCTTACGCGTCATGCTTCTTCAGTTTAAAGAAATGGTTCAGTTAGTCAAAAGATAAA	712		
Db	77	AACCTTACGCGTCATGCTTCTTCAGTTTAAAGAAATGGTTCAGTTAGTCAAAAGATAAA	136		
QY	713	ATAATAGATATTTCCATTTAAATATGATCCGAGAAAAGATTGGAGGTTTTGGCATATAGG	772		
Db	137	ATAATAGATATTTCCATTTAAATATGATCCGAGAAAAGATTGGAGGTTTTGGCATATAGG	196		
QY	773	TAAATACTAGATGATATCGAATGTGCTCAAGAAAATCCATAGCTATTTTCTACCCGAGATA	832		
Db	197	TAAATACTAGATGATATCGAATGTGCTCAAGAAAATCCATAGCTATTTTCTACCCGAGATA	256		
QY	833	CCATTCCTTAAAGGGCGGCTATTATGACAAACCACTGCAAAATATGCGATCAGACGAGTAAAG	892		
Db	257	CCATTCCTTAAAGGGCGGCTATTATGACAAACCACTGCAAAATATGCGATCAGACGAGTAAAG	316		
QY	893	AGTTCCTTGATTCATCCGCCGAATACACAAATGGGAATTTGGGGCGTTCATGCGAGTATCC	952		
Db	317	AGTTCCTTGATTCATCCGCCGAATACACAAATGGGAATTTGGGGCGTTCATGCGAGTATCC	376		
QY	953	ATTTCCTCTTAAACCGCGCATGCGATGATGATGATGATGATTTTGAAGAGATGTTGATTCAA	1012		
Db	377	ATTTCCTCTTAAACCGCGCATGCGATGATGATGATGATGATTTTGAAGAGATGTTGATTCAA	436		
QY	1013	TGAATCATCATGCTGATGCCCGTAGCAAGTTGCGTGAAGAAATTAGCTGAGCTTACCGCGG	1072		
Db	437	TGAATCATCATGCTGATGCCCGTAGCAAGTTGCGTGAAGAAATTAGCTGAGCTTACCGCGG	496		
QY	1073	AATTAAAGATTTTATTCAGTTATTCACGCCGAATTTAATATAGCATCTGCTAGTAGTGCCA	1132		
Db	497	AATTAAAGATTTTATTCAGTTATTCACGCCGAATTTAATATAGCATCTGCTAGTAGTGCCA	556		
QY	1133	CCATTAATATCCATGATATAATCCATTATTCATGATGATGATGATGATTTATATGTTATACG	1192		
Db	557	CCATTAATATCCATGATATAATCCATTATTCATGATGATGATGATGATTTATATGTTATACG	616		
QY	1193	ATGAAGATATTTTAAACCGACGACGAGATACAAAATTCCTCGAGAAAATGCGCTCAACCA	1252		
Db	617	ATGAAGATATTTTAAACCGACGACGAGATACAAAATTCCTCGAGAAAATGCGCTCAACCA	676		
QY	1253	CCATTCAGGTGATGGAGCGAGAAAAAAATATGCTCGATTAAGAGACTTCTTGGAGAGTG	1312		
Db	677	CCATTCAGGTGATGGAGCGAGAAAAAAATATGCTCGATTAAGAGACTTCTTGGAGAGTG	736		
QY	1313	AGAAATAAAAAGAACCGGGCGTGGGTAATCTGAAAAACCTCATACTCTTATAATAAGATA	1372		
Db	737	AGAAATAAAAAGAACCGGGCGTGGGTAATCTGAAAAACCTCATACTCTTATAATAAGATA	796		
QY	1373	ATAATGAAATATCTACACTTGGCACCACTGCTCGGATTAAGTCAAGCGCGCAACGACT	1432		
Db	797	ATAATGAAATATCTACACTTGGCACCACTGCTCGGATTAAGTCAAGCGCGCAACGACT	856		
QY	1433	TGTTAGGCCAAAAACAACACTCAGCTGCTCGATTTACATCAGTTTATATAGCTATTTG	1492		
Db	857	TGTTAGGCCAAAAACAACACTCAGCTGCTCGATTTACATCAGTTTATATAGCTATTTG	916		
QY	1493	AAGCACTGACCGCTTTCATTCAGAAATATGATTCAGTGTGCAAGCTGCTAGATGACA	1552		
Db	917	AAGCACTGACCGCTTTCATTCAGAAATATGATTCAGTGTGCAAGCTGCTAGATGACA	976		
QY	1553	CGTCTGGTAAATGA 1566			
Db	977	CGTCTGGTAAATGA 990			

gene
1. .981
/db.xref="taxon:632"
/gene="icrv"
1. .981
/gene="icrv"
/codon_start=1
/transl_table=1
/product="V antigen"
/protein_id="AA64077.1"
/db.xref="GI:7578515"
/translation="MIRAYBNPHFIEDLENVREPOLTGHSSVLELYOYKDKNI
DISIKYPRKDSYFANRVITDDLELRKILATFLPEDALIKGHTNOJONGIKRK
EFLESSPNTQWELFAYVHFSLTADRIDDDILKVLVDSNNHGDARSLRERLAE
TAEIKYISVQAEINKMLSSGTSINIDKMLYGTDEIRKASREYKILEK
MPQTTIYDSESEKIVSINKDLSGSENRRTGALGMLKXSYKNNKNSHPATSSDK
SRPLNDIVDSKTTQSDITSFNSAIEALNFIKXDSVWGRLLDDISGR"

BASE COUNT 334 a 175 c 196 g 276 t
ORIGIN

Query Match 49.68; Score 777; DB 2; Length 981;
Best Local Similarity 99.68; Pred. No. 0;
Matches 977; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 586 ATGATTAGAGCTTACGAGCAAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGG 645
DB 1 ATGATTAGAGCTTACGAGCAAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGG 60
OY 646 GTGGAACAACCTACTGTCATGCTTCTCAGTTTAGAAGAATTGTCAGTTAGTCAA 705
DB 61 GTGGAACAACCTACTGTCATGCTTCTCAGTTTAGAAGAATTGTCAGTTAGTCAA 120
OY 706 GATAAAATATAGATTTTCCATTAATATGATCCCAAGAAAGATTGCGAGTTTGGC 765
DB 121 GATAAAATATAGATTTTCCATTAATATGATCCCAAGAAAGATTGCGAGTTTGGC 180
OY 766 AATAGATAATTAAGATGATGATGCAATGCTCAAGAAATCTAGCTTATTTTACCC 825
DB 181 AATAGATAATTAAGATGATGATGCAATGCTCAAGAAATCTAGCTTATTTTACCC 240
OY 826 GAGATACCATCTTAAAGGGGCTATATGACAAACCACTGCAAAATGCGATCAAGCA 885
DB 241 GAGATACCATCTTAAAGGGGCTATATGACAAACCACTGCAAAATGCGATCAAGCA 300
OY 886 GTAAAGAGTCTTGAATCATCGCCGAATACACAATGGGAATGCGGGCGTTATGGCA 945
DB 301 GTAAAGAGTCTTGAATCATCGCCGAATACACAATGGGAATGCGGGCGTTATGGCA 360
OY 946 GTAATGCAATTTCTTTAACCGCGATGATGATGATGATTTTGAAGTATGTT 1005
DB 361 GTAATGCAATTTCTTTAACCGCGATGATGATGATGATTTTGAAGTATGTT 420
OY 1006 GATTCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
DB 421 GATTCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
OY 1066 ACCGCGAATTAAGATTTATTCAGTATTAAGCGGAATTAATAGCATCTGCTAGT 1125
DB 481 ACCGCGAATTAAGATTTATTCAGTATTAAGCGGAATTAATAGCATCTGCTAGT 540
OY 1126 AGTGGACACATTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
DB 541 AGTGGACACATTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
OY 1186 TATACAGTGAAGATTTTAAAGCCAGCGAGTACAAAATTCGAGAAAAATGCGCT 1245
DB 601 TATACAGTGAAGATTTTAAAGCCAGCGAGTACAAAATTCGAGAAAAATGCGCT 660
OY 1246 CAACACCATTCAGTGGATGGAGGAGAGAAAAAATAGTGTGATAAGGACTTCTT 1305
DB 661 CAACACCATTCAGTGGATGGAGGAGAGAAAAAATAGTGTGATAAGGACTTCTT 720
OY 1306 GGAAGTGAATAAAGAACGGGGCGTTGGGTAATCTGAAAAACTCATCTTATTAAT 1365

DB 721 GGAAGTGAATAAAGAACGGGGCGTTGGGTAATCTGAAAACTCATCTTATTAAT 780
OY 1366 AAGATATAATGAATTAATCATCTTGGCCACACCGCTGCGATTAAGTCCAGCCGCTC 1425
DB 781 AAGATATAATGAATTAATCATCTTGGCCACACCGCTGCGATTAAGTCCAGCCGCTC 840
OY 1426 AACGACTTGGTTAGCCAAAACCAACACGCTGCTGATATTAATGATGATGATGATGATGAT 1485
DB 841 AACGACTTGGTTAGCCAAAACCAACACGCTGCTGATATTAATGATGATGATGATGATGAT 900
OY 1486 GCTATTGAGACGCTGAGAACCGTTTCATTCAGAAATATGATGATGATGATGATGATGAT 1545
DB 901 GCTATTGAGACGCTGAGAACCGTTTCATTCAGAAATATGATGATGATGATGATGATGAT 960
OY 1546 GATGACACGCTGCTGATTAATGA 1566
DB 961 GATGACACGCTGCTGATTAATGA 981

RESULT 13
A56812
LOCUS A56812 547 bp DNA PAT 03-MAR-1998
DEFINITION Sequence 20 from Patent WO9628551.
ACCESSION A56812
VERSION A56812.1 GI:3712825
KEYWORDS
SOURCE
ORGANISM
Yersinia pestis.
Bacteria; Proteobacteria; gamma subphylum; Enterobacteriaceae;
Yersinia.
REFERENCE
AUTHORS 1 (bases 1 to 547)
TITLE T1bball R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C.,
Bennett and Alice, M.
JOURNAL VACCINES FOR PLAGUE
COMMENT Patent: WO 9628551-A 20 19-SEP-1996;
SECUR DEFENCE (GB)
Other publication ZA 9602036 960716
Other publication AU 4951196 961002.
FEATURES
source
location/Qualifiers
1..547
/organism="Yersinia pestis"
/db.xref="taxon:632"
29..>538
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA03422.1"
/translation="MKRISVIAIALEGTATNADLTASTATATATVAPRTLLTY
KEBAPTINMNGINDIELYGLTGLGKRGTTSTSVNFTDAAGDPRLYFTSDGDN
HOFITKVGDSDFDISFRVNEENLVATGSDVLTGSDVFTSISGSKLAKRTT
DAVTATVSNQ"

BASE COUNT 165 a 120 c 115 g 147 t
ORIGIN

Query Match 32.88; Score 513; DB 5; Length 547;
Best Local Similarity 100.0%; Pred. No. 2e-247;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 ATATGAAAAAATAGTCCGTATTCGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 127
DB 27 ATATGAAAAAATAGTCCGTATTCGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 86
OY 128 ATGGGCGAGATTAACTGCAAGCAACCACTGCAAGCGCACTCTGTTGAAACAGCCCGCA 187
DB 87 ATGGGCGAGATTAACTGCAAGCAACCACTGCAAGCGCACTCTGTTGAAACAGCCCGCA 146
OY 188 TCACCTTTACATATAGAGAGCGGCTCCCAATTAATGATGATGATGATGATGATGATGATGAT 247
DB 147 TCACCTTTACATATAGAGAGCGGCTCCCAATTAATGATGATGATGATGATGATGATGATGAT 206

OY	248	CAGATTACTCTGTGGTACGCTTACTCTTGCGGGCATTAATAACAGAACCACTATCACAT	307
Db	207	CAGAATTACTCTGTGGTAGCCTTACTCTTTGGCGGCTATTAACAGAACCACTATCACAT	266
OY	308	CTGTTAACTTTAAGATGCCGGGGGTGATCCCATGTACTTAACAATTACTTCTCAGATG	367
Db	267	CTGTTAACTTTACAGATGGCGGGGTGATCCCATGTACTTAACAATTACTTCTCAGATG	326
OY	368	GAAATACCACCAATTCACATAAAGATGTGGCAAGATTTCTAGATTTTTGATACCT	427
Db	327	GAAATACCACCAATTCACATAAAGATGTGGCAAGATTTCTAGATTTTTGATACCT	386
OY	428	CTCCTAAGGTAAACGGTGAACACTTGTGGGGATGACAGCTGCTTGGCTACGGGCAGCC	487
Db	387	CTCCTAAGGTAAACGGTGAACACTTGTGGGGATGACAGCTGCTTGGCTACGGGCAGCC	446
OY	488	AGGATTTCTTTGTGCTCATATTTGGTCTCCAAGGGGGTAAACTTCAGACAGGTAAATCA	547
Db	447	AGGATTTCTTTGTGCTCATATTTGGTCTCCAAGGGGGTAAACTTCAGACAGGTAAATCA	506
OY	548	CTGATGCTGTACCGTAAACCGTATCTAACCAAG	580
Db	507	CTGATGCTGTACCGTAAACCGTATCTAACCAAG	539
RESULT	14		
YPCAF			
LOCUS	5383 bp	DNA	BCT
DEFINITION	Y. pestis genes cafI, cafIM, cafIIA and cafIR.		23-NOV-1995
ACCESSION	X61996.S40525 S90405 X57773		
VERSION	X61996.1 GI:48620		
KEYWORDS	cafI gene; cafIIA gene; cafIM gene; cafIR gene; fI antigen.		
SOURCE	Yersinia pestis.		
ORGANISM	Yersinia pestis		
REFERENCE	Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.		
AUTHORS	1 (bases 1 to 5383) Gal'ov,E.E., Smirnov,O.Yu., Karlishhev,A.V., Chernovskaya,T.V., Dolgikh,D.A., Smirnov,O.Yu., Volkovoy,K.I., Abramov,V.M. and Zav'yalov,V.P. Expression of the envelope antigen fI of Yersinia pestis is mediated by the product of cafim gene having homology with the chaperone protein Pamp of Escherichia coli		
TITLE	FEMS Lett. 286 (1-2), 79-82 (1991)		
JOURNAL	91323540		
MEDLINE	2 (bases 1 to 5383)		
REFERENCE	Gal'ov,E.E., Smirnov,O.Yu., Karlishhev,A.V., Volkovoy,K.I., Denesyuk,A.I., Nazimov,I.V., Rubtsov,K.S., Abramov,V.M., Daladyanz,S.M. and Zav'yalov,V.P. Nucleotide sequence of the Yersinia pestis gene encoding fI antigen and the primary structure of the protein. Putative T and B cell epitopes		
AUTHORS	FEMS Lett. 277 (1-2), 230-232 (1990)		
TITLE	91099503		
JOURNAL	3 (bases 1 to 5383)		
MEDLINE	Karlishhev A.V.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (24-JAN-1992) A.V. Karlishhev, Inst of Immunology, State Concern 'Biopreparation', 142380 Lyubuchany, Moscow Region, USSR revised by [5]		
TITLE	4 (bases 1 to 5383)		
JOURNAL	Karlishhev,A.V., Gal'ov,E.E., Abramov,V.M. and Zav'yalov,V.P.		
MEDLINE	CafI gene and its role in the regulation of capsule formation of Y. pestis		
REFERENCE	FEMS Lett. 305 (1), 37-40 (1992)		
AUTHORS	5 (bases 1 to 5383)		
TITLE	Karlishhev,A.V.		
JOURNAL	Direct Submission		
MEDLINE	Submitted (22-SEP-1992) to the EMBL/Genbank/DBJ databases		
REFERENCE	6 (bases 1 to 5383)		
AUTHORS	Karlishhev,A.V., Gal'ov,E.E., Smirnov,O.Yu., Guzeyev,A.P., Abramov,V.M. and Zav'yalov,V.P.		
TITLE	A new gene of the fI operon of Y. pestis involved in the capsule		

bioRxiv preprint doi: <https://doi.org/10.1101/1992>; this version posted September 29, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Wed Aug 23 11:45:50 2000

us-08-699-716a-1.oligo.rge

Page 25

Search completed: August 22, 2000, 16:32:27
Job time: 4840 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 15:39:08 ; Search time 63.88 Seconds
(without alignments)
6133.388 Million cell updates/sec

Title: US-08-699-716a-1

Perfect score: 1566
Sequence: 1 ATGGGCATCATCATCATCA.....ATGACACGTCTGTGAATGA 1566

Scoring table: OLIGO_NMC
Gapop 60.0, Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size: 0

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database: N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	927	59.2	1014	1 T04222	Partial LcrV (V an
2	923	58.9	1014	1 T38243	Y. pestis lcrV (V
3	876	55.9	1014	1 T38242	Y. pestis lcrV (V
4	876	55.9	1462	1 T38256	Y. pestis Fl/V ant
5	876	55.9	1530	1 T38249	Y. pestis Fl/V ant
6	872	55.7	1014	1 T04223	Partial LcrV (V an
7	513	32.8	547	1 T38248	Y. pestis Fl antiq
8	512	32.7	542	1 Q92819	Yersinia pestis ca
9	512	32.7	544	1 V41594	Nucleotide sequenc
10	512	32.7	544	1 V41596	Nucleotide sequenc
11	510	32.6	510	1 V41595	Nucleotide sequenc
12	448	28.6	450	1 V41601	Nucleotide sequenc
13	448	28.6	474	1 V41600	Nucleotide sequenc
14	447	28.5	447	1 V41609	Nucleotide sequenc
15	402	25.7	576	1 V41598	Nucleotide sequenc
16	398	25.4	541	1 Q92817	Yersinia pestis ca
17	398	25.4	541	1 T38244	Y. pestis catI (Fl
18	396	25.3	542	1 Q92818	Yersinia pestis ca
19	339	21.6	513	1 V41599	Nucleotide sequenc
20	68	4.3	375	1 V41599	Nucleotide sequenc
21	68	4.3	375	1 V41599	Nucleotide sequenc
22	68	4.3	375	1 V41599	Nucleotide sequenc
23	68	4.3	375	1 V41599	Nucleotide sequenc
24	68	4.3	375	1 V41599	Nucleotide sequenc
25	68	4.3	375	1 V41599	Nucleotide sequenc
26	68	4.3	375	1 V41599	Nucleotide sequenc
27	68	4.3	375	1 V41599	Nucleotide sequenc
28	68	4.3	375	1 V41599	Nucleotide sequenc
29	68	4.3	375	1 V41599	Nucleotide sequenc
30	68	4.3	375	1 V41599	Nucleotide sequenc
31	68	4.3	375	1 V41599	Nucleotide sequenc
32	68	4.3	375	1 V41599	Nucleotide sequenc
33	68	4.3	375	1 V41599	Nucleotide sequenc

34	53	3.4	1547	1 V30581	Clostridium botuli
35	35	2.2	47	1 T37132	Sense primer T20-I
36	36	2.2	861	1 O81500	SVF anti-rev seque
37	35	2.2	861	1 T45347	Single chain SVF a
38	30	1.9	416	1 O59427	Human brain expres
39	30	1.9	8920	1 O62924	Cardamoyl-phosphat
40	29	1.9	257	1 V88713	EST clone RH378. N
41	29	1.9	598	1 N90919	Type A insertion g
42	29	1.9	1021	1 T04613	5' flanking region
43	29	1.9	2359	1 Q58609	Sequence encoding
44	29	1.9	2787	1 X15661	Protein phosphatas
45	29	1.9	4459	1 V99094	DNA methyltransfer
46	29	1.9	4641	1 N81538	Sequence of a gene
47	29	1.9	19440	1 N99129	DNA methyltransfer
48	28	1.8	39	1 T05194	Human/mouse quanyl
49	28	1.8	39	1 X15374	Human quanylate ki
50	28	1.8	76	1 V08564	Primer for Transge
51	28	1.8	86	1 T11965	Plasmid pET15b His
52	28	1.8	287	1 V13947	Salmonella sefa ge
53	28	1.8	552	1 X29725	Insert from pET28C
54	28	1.8	1247	1 V08559	Transgene for epit
55	28	1.8	1310	1 V08558	Topoisomerase I ex
56	28	1.8	2136	1 V30788	Transgene for epit
57	28	1.8	3263	1 V08560	Transgene for epit
58	28	1.8	4286	1 V08559	Nucleotide sequenc
59	27	1.7	38	1 V41618	Human gene signalu
60	27	1.7	250	1 T22691	DNA primer for Yer
61	25	1.6	25	1 Q92816	Y. pestis Fl antiq
62	25	1.6	25	1 T38247	Viral infection ge
63	25	1.6	904	1 V06104	Human gene signalu
64	24	1.5	230	1 T24068	Recombinant mite a
65	24	1.5	447	1 Q92652	Alcohol oxidase ge
66	24	1.5	1667	1 Q48589	Nucleotide sequenc
67	23	1.5	30	1 V41617	DNA primer for Yer
68	23	1.5	32	1 Q92814	Y. pestis Fl antiq
69	23	1.5	32	1 T38245	Cmb antisense PCR
70	23	1.5	47	1 V58140	Mycobacterium anti
71	23	1.5	53	1 V44450	Mycobacterium anti
72	23	1.5	53	1 V44407	M. tuberculosis fu
73	23	1.5	53	1 V64559	M. tuberculosis fu
74	23	1.5	53	1 V64516	Human gene signalu
75	23	1.5	135	1 T25107	DNA encoding a hum
76	23	1.5	553	1 X51720	cayAEI DNA sequenc
77	23	1.5	797	1 V36459	Mouse llyor-1 gen
78	23	1.5	5434	1 V64410	Sequence encoding
79	23	1.5	7635	1 Q37811	Human neuronal cal
80	23	1.5	7635	1 O84654	DNA encoding human
81	23	1.5	7635	1 V42679	DNA encoding human
82	23	1.5	7635	1 V42679	Mycobacterium anti
83	23	1.5	7635	1 V42679	Nucleotide sequenc
84	23	1.5	7676	1 V55801	Nucleotide sequenc
85	23	1.5	7676	1 V64567	Nucleotide sequenc
86	22	1.4	33	1 V41608	Amp antisense PCR
87	22	1.4	45	1 V58142	Expression plasmid
88	22	1.4	1020	1 N92873	Expression plasmid
89	22	1.4	1518	1 N92872	Expression plasmid
90	22	1.4	1665	1 N92871	Expression plasmid
91	22	1.4	2079	1 N94527	Peptidyl C-termina
92	22	1.4	2625	1 O87970	cDNA of plasmid px
93	22	1.4	3315	1 N90791	Nucleotide sequenc
94	21	1.3	30	1 V41616	Nucleotide sequenc
95	21	1.3	37	1 V41616	Reverse primer for
96	21	1.3	47	1 T72159	Primer for amplify
97	21	1.3	97	1 T29649	Ich-2 gene 5' prim
98	21	1.3	104	1 T33359	EST clone CT720. N
99	21	1.3	262	1 V89773	Murine C140 recept
100	21	1.3	1477	1 O84557	Murine C140 recept
101	21	1.3	1477	1 T32036	Sequence of genom
102	21	1.3	1497	1 T47066	pinto bean alpha-D
103	21	1.3	1648	1 Q25532	Murine C140 recept
104	21	1.3	2732	1 O84559	Sequence of genom
105	21	1.3	2732	1 T32038	Murine C140 recept
106	20	1.3	20	1 O55602	5' flanking sequen

Clostridium botuli
Sense primer T20-I
SVF anti-rev seque
Single chain SVF a
Human brain expres
Cardamoyl-phosphat
EST clone RH378. N
Type A insertion g
5' flanking region
Sequence encoding
Protein phosphatas
DNA methyltransfer
Sequence of a gene
DNA methyltransfer
Human/mouse quanyl
Human quanylate ki
Primer for Transge
Plasmid pET15b His
pET-15b expression
Salmonella sefa ge
Insert from pET28C
Transgene for epit
Topoisomerase I ex
Transgene for epit
Transgene for epit
Nucleotide sequenc
Human gene signalu
DNA primer for Yer
Y. pestis Fl antiq
Viral infection ge
Human gene signalu
Recombinant mite a
Alcohol oxidase ge
Nucleotide sequenc
DNA primer for Yer
Y. pestis Fl antiq
Cmb antisense PCR
Mycobacterium anti
Mycobacterium anti
M. tuberculosis fu
M. tuberculosis fu
Human gene signalu
DNA encoding a hum
cayAEI DNA sequenc
Mouse llyor-1 gen
Sequence encoding
Human neuronal cal
Human neuronal cal
DNA encoding human
DNA encoding human
Mycobacterium anti
M. tuberculosis fu
Nucleotide sequenc
Amp antisense PCR
Expression plasmid
Expression plasmid
Expression plasmid
Peptidyl C-termina
cDNA of plasmid px
Nucleotide sequenc
Nucleotide sequenc
Reverse primer for
Primer for amplify
Ich-2 gene 5' prim
EST clone CT720. N
Murine C140 recept
Murine C140 recept
Sequence of genom
pinto bean alpha-D
Murine C140 recept
Sequence of genom
Murine C140 recept
5' flanking sequen

c 399	16	1.0	4004	T51260	Human AD4 gene gen	c 472	15	1.0	384	1	T29972	Human gene signatu
c 400	16	1.0	4259	V04201	Merged contigs pfr	c 473	15	1.0	399	1	T44935	Partial sequence o
c 401	16	1.0	4317	V52257	Streptococcus pneu	c 474	15	1.0	419	1	T30731	Mouse cryptidin 6 g
c 402	16	1.0	4464	N71244	Genomic sequence o	c 475	15	1.0	422	1	T30740	Mouse cryptidin 2 c
c 403	16	1.0	4465	Q28697	Glutamine syntheta	c 476	15	1.0	422	1	T30744	Mouse cryptidin 6 c
c 404	16	1.0	5663	X12989	Enterococcus faeca	c 477	15	1.0	435	1	V70865	Internal transcrib
c 405	16	1.0	6008	Q41289	Ubiquitin-specific	c 478	15	1.0	426	1	V75379	Streptococcus aur
c 406	16	1.0	6516	V52238	Streptococcus pneu	c 479	15	1.0	445	1	T13906	Human gene signatu
c 407	16	1.0	6619	V63789	Plasmid pKM72/68 e	c 480	15	1.0	468	1	X21097	Polynucleotide seq
c 408	16	1.0	6761	X20517	Polynucleotide seq	c 481	15	1.0	478	1	T83960	DNA encoding a Sta
c 409	16	1.0	7296	X12996	Enterococcus faeca	c 482	15	1.0	533	1	Q27947	G-CSF receptor ago
c 410	16	1.0	8133	T29248	C. difficile toxin	c 483	15	1.0	534	1	T66430	B.campestris seed-
c 411	16	1.0	8133	V30560	Clostridium diffic	c 484	15	1.0	540	1	Q13860	CDNA clone pms14 e
c 412	16	1.0	10023	V52731	Human hepatocyte n	c 485	15	1.0	581	1	Q05760	CDNA clone pms14 e
c 413	16	1.0	10897	T09187	Mutu putative onco	c 486	15	1.0	581	1	Q08573	CDNA clone pms14 e
c 414	16	1.0	11191	X20578	Polynucleotide seq	c 487	15	1.0	581	1	Q08545	CDNA clone pms14 e
c 415	16	1.0	12011	V83540	PCR-generated regl	c 488	15	1.0	581	1	T60953	Tapetum-specific c
c 416	16	1.0	13542	V74465	Staphylococcus aur	c 489	15	1.0	581	1	Q05668	pms14 contig. male
c 417	16	1.0	14078	V74502	Staphylococcus aur	c 490	15	1.0	608	1	V17547	Peptidyl prollyl ci
c 418	16	1.0	14516	X06748	E. coli O111 antiq	c 491	15	1.0	614	1	Q34547	IL-6 sss muten.
c 419	16	1.0	15462	V18272	HP1V-3 JS isolate	c 492	15	1.0	614	1	Q08361	Human interleukin-
c 420	16	1.0	15462	V18273	HP1V-3 FRh1 cp45 v	c 493	15	1.0	614	1	T33895	Coding sequence fo
c 421	16	1.0	15462	V18274	HP1V-3 Vero cp45 v	c 494	15	1.0	615	1	T67398	H. pylori cytoplas
c 422	16	1.0	15462	V70401	Human parainfluenz	c 495	15	1.0	633	1	V72021	Adenovirus 15S1gDE
c 423	16	1.0	15462	V83561	Nucleotide sequenc	c 496	15	1.0	645	1	V74997	Staphylococcus aur
c 424	16	1.0	15660	V83532	Plasmid p3/7(131)	c 497	15	1.0	658	1	T33735	Enterococcus faeca
c 425	16	1.0	15666	V83533	Plasmid p3/7(131)2	c 498	15	1.0	686	1	X16869	DNA encoding a S.
c 426	16	1.0	15669	V83520	Plasmid p218(131)	c 499	15	1.0	686	1	X14334	H. pylori GHPO 212
c 427	16	1.0	19718	V52232	Streptococcus pneu	c 500	15	1.0	692	1	V74448	Staphylococcus aur
c 428	16	1.0	35524	V22140	Chimpanzee adenovi	c 501	15	1.0	709	1	X30726	Streptococcus pneu
c 429	16	1.0	45546	X23520	Human kidney amino	c 502	15	1.0	717	1	T63298	Human proteasome a
c 430	16	1.0	110000	V21209_00	Methanococcus jam	c 503	15	1.0	777	1	X14447	H. pylori GHPO 902
c 431	16	1.0	110000	V21209_10	Continuation (11 o	c 504	15	1.0	781	1	V06144	Viral infection ge
c 432	16	1.0	133894	T13893	ACNPV genomic DNA	c 505	15	1.0	786	1	T13721	ACNPV ORF 141, res
c 433	15	1.0	19	V35789	Human epimorphin c	c 506	15					

C	983	14	0.9	513	1	V41599	Nucleotide sequenc
	984	14	0.9	514	1	V02153	Human secreted pro
	985	14	0.9	514	1	T88067	Partial cDNA clone
C	986	14	0.9	514	1	V07572	Human mago nashi p
	987	14	0.9	516	1	V26396	Vascular endothell
	988	14	0.9	519	1	V24797	H. pylori ORF 11ge
	989	14	0.9	519	1	V24509	H. pylori cellular
	990	14	0.9	522	1	T64615	G-CSF receptor ago
	991	14	0.9	522	1	T64616	G-CSF receptor ago
	992	14	0.9	522	1	T64617	G-CSF receptor ago
	993	14	0.9	532	1	T64618	G-CSF receptor ago
	994	14	0.9	535	1	Q04481	Plasmid pAS28 enco
	995	14	0.9	535	1	Q04482	Plasmid pASN6 enco
	996	14	0.9	535	1	Q04484	Plasmid pASN145 en
	997	14	0.9	535	1	N80947	G-CSF gene isolate
	998	14	0.9	535	1	O87151	Human granulocyte
	999	14	0.9	535	1	T64619	G-CSF receptor ago
	1000	14	0.9	535	1	T64620	G-CSF receptor ago

ALIGNMENTS

RESULT	1
ID	T04222
AC	T04222; standard; DNA; 1014 BP.
DT	18-APR-1996 (first entry)
DE	Partial LcrV (V antigen) gene of <i>Y. pestis</i> .
LCRV	V antigen; virulence; plague; vaccine; eptlopte; ss.
OS	<i>Yersinia pestis</i> .
FT	Location/Qualifiers
FT	1..990
FT	/*tag= a
FT	/note= "V antigen"
PN	MO9524475-A1.
PD	14-SEP-1995.
PF	06-MAR-1995; G00481.
PR	08-MAR-1994; GB-004577.
PA	(M1NA) UK SEC FOR DEFENCE.
PI	Leary SEC, Tibbital RM, Williamson ED, Leary SE;
PI	WPI; 95-328268//42.
DR	P-RSDS: R79961.
PT	Recombinant DNA expressing <i>Yersinia pestis</i> V antigen - useful in
PT	oral or parenteral vaccines for protection against plague
PS	Claim 6; Page 11-13; 25pp; English.
PS	T04222-23 are DNA sequences (lcrV) encoding all or a protective eptloptic
CC	part of the mature V protein of <i>Yersinia pestis</i> . The protein was
CC	expressed as a fusion protein with maltose binding protein or
CC	glutathione-S-transferase in 3 different plasmid vectors. <i>Y. pestis</i> is
CC	the highly virulent causative organism of plague in a wide range of
CC	animals, including man. The V antigen (lcrV) is an unstable 37.3 kDa
CC	monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen
CC	is postulated to act as a virulence antigen, and transformed
CC	microorganisms contg. recombinant DNA encoding a V antigen protein/
CC	peptide are useful in vaccines to protect against plague.
Q	Sequence 1014 BP: 346 A: 181 C: 201 G: 286 T;

Query Match	59.2%	Score 927	DB 1	Length 1014
Best Local Similarity	99.9%	Pred. No. 0		
Matches 977	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

QY	ATTGAGACCTCGAACAACCAACCAATTTTATGAGATCTAGAAAAAGTTAGGTC	648
QY	589 ATTGAGACCTCGAACAACCAACCAATTTTATGAGATCTAGAAAAAGTTAGGTC	
Db	13 ATTGAGACCTCGAACAACCAACCAATTTTATGAGATCTAGAAAAAGTTAGGTC	72
QY	649 GAAACACTTACTGTCATGTTCTTCAGTTTAGAAGATTTGGTTCAGTATGCAAGAT	708
Db	73 GAAACACTTACTGTCATGTTCTTCAGTTTAGAAGATTTGGTTCAGTATGCAAGAT	1322
QY	709 AAAAATATGATATTTCATTAATATGATCCGAAAAAGTTGGAGTTTGGCAAT	768

D	b	133	AAAATAATAGATTTCACCTTAATAATGANTCCAGAAAAAGATTGGAGGTTTTCCAA	T	192
O	y	769	AGAGTAATTAAGTAGATGATATGCAGTAATGGCTCAGAGAAAATCCTAGCTTAATTTCTACCAG	A	828
D	b	193	AGAATAATTAAGTAGATGATATGCAGTAATGGCTCAGAGAAAATCCTAGCTTAATTTCTACCAG	A	252
O	y	829	GATACCAATCTTAAAGCGGCTCATTTTGACAACCAACTGCAAAATGGCATCAAGCGA	T	888
D	b	253	GATGCCATCTTAAAGCGGCTCATTTTGACAACCAACTGCAAAATGGCATCAAGCGA	T	312
O	y	889	AAAGAGTTCCTTGATCATCGCCGAATPACAAATGGGAATTCGGGGCTTCATGGCAG	T	948
D	b	313	AAAGAGTTCCTTGATCATCGCCGAATPACAAATGGGAATTCGGGGCTTCATGGCAG	T	372
O	y	949	ATGATTTCTCTTAACCGCGGAGTCCATGCGATGATGATATTTTAAAGTATTTGAT	T	1008
D	b	373	ATGATTTCTCTTAACCGCGGAGTCCATGCGATGATGATATTTTAAAGTATTTGAT	T	432
O	y	1009	TCAATGATCATCATCGTGTGATGCCCGTAGCAAGTTGCTGGAAGATTAAGTAGCTAAC	C	1068
D	b	433	TCAATGATCATCATCGTGTGATGCCCGTAGCAAGTTGCTGGAAGATTAAGTAGCTAAC	C	492
O	y	1069	GCCGAATTTAAAGATTATTCAGTTATTCAGCCGGAATTAATPAAGCATCTGTAGT	A	1128
D	b	493	GCCGAATTTAAAGATTATTCAGTTATTCAGCCGGAATTAATPAAGCATCTGTAGT	A	552
O	y	1129	GGCACCAATTAATATCCATGATTAATATCCATTAATCTCATGATGATTAATTAATGAT	T	1188
D	b	553	GGCACCAATTAATATCCATGATTAATATCCATTAATCTCATGATGATTAATTAATGAT	T	612
O	y	1189	ACAGATGAGAGATTTTTTAAAGCAGCGCAGATACAAATTTCTGAGAAATGGCTCA	A	1248
D	b	613	ACAGATGAGAGATTTTTTAAAGCAGCGCAGATACAAATTTCTGAGAAATGGCTCA	A	672
O	y	1249	ACCACCATTCAGGTGGATGGGAGCGAGAAAAATATGCTCGATTAAGSACTTCTTGA	A	1308
D	b	673	ACCACCATTCAGGTGGATGGGAGCGAGAAAAATATGCTCGATTAAGSACTTCTTGA	A	732
O	y	1309	AGTAGAGATTAAGAACCGGGGGCTTGGGTAATCTGAANAATCTACTACTTATATAA	A	1368
D	b	733	AGTAGAGATTAAGAACCGGGGGCTTGGGTAATCTGAANAATCTACTACTTATATAA	A	792
O	y	1369	GATAATAATGATTAATCTCACTTGGCACCACTGCTCGATTAAGTCAGGCCGCTCAAC	A	1428
D	b	793	GATAATAATGATTAATCTCACTTGGCACCACTGCTCGATTAAGTCAGGCCGCTCAAC	A	852
O	y	1429	GACTTGTTAGCCAAAAAACAACCTCAGCTGCTGATATTCACACGTTTAACTCACT	A	1488
D	b	853	GACTTGTTAGCCAAAAAACAACCTCAGCTGCTGATATTCACACGTTTAACTCACT	A	912
O	y	1489	ATTGAGACACTGAACCGTTTCATTCAGAAATATGATTCAGATCAACGTCGTAGAT	A	1548
D	b	913	ATTGAGACACTGAACCGTTTCATTCAGAAATATGATTCAGATCAACGTCGTAGAT	A	972
O	y	1549	GACACGTCGTGTAATGA	A	1566
D	b	973	GACACGTCGTGTAATGA	A	990

RESULT	2
ID	T38243
AC	T38243 standard; DNA; 1014 BP.
DT	28-DEC-1996. (first entry)
DE	Y. pestis lcrV (V antigen) gene.
KW	Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW	Fl antigen; ds.
OS	Yersinia pestis strain GB.
FT	Key
FT	Location/Qualifiers
FT	1..990
FT	/*tag- a
FT	1..10
FT	/*tag- b
FT	misc-feature

OY	1189	ACAGATGGAAGATTCTTTTAAAGCCAGCGAGGTACAAAATTCGCAGAATAATGCCCTCAA	1248
Dd	1141	ACAGATGGAAGATTCTTTTAAAGCCAGCGAGGTACAAAATTCGCAGAATAATGCCCTCAA	1200
OY	1249	ACCACCATTCAGGTGGATGGAGCGGAGAAAAAATAATAGTCTCGATAAAGACTTCTTGGA	1308
Dd	1201	ACCACCATTCAGGTGGATGGAGCGGAGAAAAAATAATAGTCTCGATAAAGACTTCTTGGA	1260
OY	1309	AGTGGAATTAAGAACAACCGGGCGTTGGGTAAATCTGAAAAACCTCATTCTATATAATAA	1368
Dd	1261	AGTGGAATTAAGAACAACCGGGCGTTGGGTAAATCTGAAAAACCTCATTCTATATAATAA	1320
OY	1369	GATAATTAATGAATTTCTCTCACTTTGGCCACCACCTGCTCGATAAATGCAAGCGCGTCAAC	1428
Dd	1321	GATAATTAATGAATTTCTCTCACTTTGGCCACCACCTGCTCGATAAATGCAAGCGCGTCAAC	1380
OY	1429	CACCTGGTGTAGCCAAAAACAACCTACGCTGTCTGATATTATCATCACGTTTTAATTCAGCT	1488
Dd	1381	GACCTGGTGTAGCCAAAAACAACCTACGCTGTCTGATATTATCATCACGTTTTAATTCAGCT	1440
OY	1489	ATTGGAAGCCTCAACCGTTTCATTTCGAAATTAATGATTCAGTAGTATGCAAGCTCGTAGAT	1548
Dd	1441	ATTGGAAGCCTCAACCGTTTCATTTCGAAATTAATGATTCAGTAGTATGCAAGCTCGTAGAT	1500
OY	1549	GACACGCTCTGTAAATGA	1566
Dd	1501	GACACGCTCTGTAAATGA	1518
RESULT	6		
T04223	ID	T04223 standard; DNA; 1014 BP.	
AC	T04223;		
DE	18-APR-1996	(first entry)	
PT	Partial LcrV (V antigen) gene of Y. pestis.		
KM	LcrV; V antigen; virulence; plague; vaccine; epitope; ss.		
OS	Yersinia pests.		
FH	Key	Location/Qualifiers	
FT	cds	1..990	
PD	W09524475-A1.	/tag= a	
PF	14-SEP-1985.		
PR	06-MAR-1985; G00481.		
PA	(MINA) UK SEC FOR DEFENCE.		
PI	Leary SEC, Titball RW, Williamson ED, Leary SE;		
DR	WPI: 95-328268/42.		
PT	R-PSDB: R79962.		
PS	Recombinant DNA expressing Yersinia pests V antigen - useful in		
CC	oral or parenteral vaccines for protection against plague		
CC	Claim 6; Page 15-16; 25pp; English.		
CC	T04222-23 are DNA sequences (lcrV) encoding all or a protective epitopic		
CC	part of the mature V protein of Yersinia pests. The protein was		
CC	expressed as a fusion protein with maltose binding protein or		
CC	glutathione-S-transferase in 3 different plasmid vectors. Y. pests		
CC	is the highly virulent causative organism of plague in a wide range of		
CC	animals, including man. The V antigen (LcrV) is an unstable 37.3 kDa		
CC	monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen		
CC	was postulated to act as a virulence antigen, and transformed		
CC	microorganisms contg. recombinant DNA encoding a V antigen protein/		
CC	peptide are useful in vaccines to protect against plague.		
CC	Sequence 1014 BP; 343 A; 185 C; 204 G; 282 T;		

Query Match	55.78;	Score 872;	DB 1;	Length 1014;
Best Local Similarity	99.88;	Pred. No. 0;		
Matches 972;	Conservative	0;	Mismatches 2;	Indels 0;
			Gaps	0

QY	553 GAGCCTTAGAACAAACCCACACAACATTTATTGAGGATCTGGAAAAAGTTGGGTGGAC 65
Dd	17 GAGCCTTAGGAACAAACCACACAACATTTATTGAGGATCTGGAAAAAGTTGGGTGGAC 76
QY	653 AACCTTACAGCGCATGCGTCTTCAGTTTATAGAAGATTTGGTTCAGTTAGTCAAAAGATATAA 712

Db	77	AACCTACGGCAGCGTCTCTTCAGTTTGAGAAATGGTTACGTTAGCAAAAGTATAAA	136
Oy	713	ATATAGATATTTCCATTAAATATGATCCAGAAAAGATTGAGAGGTTTTGGCAATAGAG	772
Db	137	ATATAGATATTTCCATTAAATATGATCCAGAAAAGATTGAGAGGTTTTGGCAATAGAG	196
Oy	773	TAATTACGATGATGATGATGATGCTCAGAAAATCCTACTTATTTTGACACCGGAGATA	832
Db	197	TAATTACTGATGATGATGATGCTCAGAAAATCCTACTTATTTTGACACCGGAGATG	256
Oy	833	CCATTCTTAAAGGGCGTCATTATGACAACTGCAAAATGGCATCAACGAGATPAAAG	892
Db	257	CCATTCTTAAAGGGCGTCATTATGACAACTGCAAAATGGCATCAACGAGATPAAAG	316
Oy	893	AGTTCCTTGAATCATGCGCGAATACACAAATGGGAATGGCGGCGTTCATGGCAGTAATGC	952
Db	317	AGTTCCTTGAATCATGCGCGAATACACAAATGGGAATGGCGGCGTTCATGGCAGTAATGC	376
Oy	953	ATTTCCTTTAACCGCGCGATCGATGAGATGATATTTTGAAGAAGATGTTGATTTCA	1012
Db	377	ATTTCCTTTAACCGCGCGATCGATGAGATGATATTTTGAAGAAGATGTTGATTTCA	436
Oy	1013	TGAATCATCATGCTGATGCCCGTAGCAAGTTGGCGTAGAATTAAGTCACTGAGCTTACCGCG	1072
Db	437	TGAATCATCATGCTGATGCCCGTAGCAAGTTGGCGTAGAATTAAGTCACTGAGCTTACCGCG	496
Oy	1073	AATTAAGATTTATTCAGTTATTCACAGCCGAATTAATAGATCTGCTAGTAAATGGCA	1132
Db	497	AATTAAGATTTATTCAGTTATTCACAGCCGAATTAATAGATCTGCTAGTAAATGGCA	556
Oy	1133	CCATTAATATCCAGATTAATCCATTATCTCATGATPAAATTTATATGTTATACAG	1192
Db	557	CCATTAATATCCAGATTAATCCATTATCTCATGATPAAATTTATATGTTATACAG	616
Oy	1193	ATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAAATTCGAGAAAATGCTCTAAACCA	1252
Db	617	ATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAAATTCGAGAAAATGCTCTAAACCA	676
Oy	1253	CCATTACAGGTGGAGTGGGAGGAGAAAAAATAGTCTGATAAGAACTTCTTGGAACTG	1312
Db	677	CCATTACAGGTGGAGTGGGAGGAGAAAAAATAGTCTGATAAGAACTTCTTGGAACTG	736
Oy	1313	AGAAATAAAGAACCCGGGCGTTGGGTATCTGAAAAACTCATACTCTTATAATAAAGATA	1372
Db	737	AGAAATAAAGAACCCGGGCGTTGGGTATCTGAAAAACTCATACTCTTATAATAAAGATA	796
Oy	1373	ATAATGATTTATCTCACTTTGGCCACCACTGCTCGGATPAGTCCAGGCGGCTCAACGACT	1432
Db	797	ATAATGATTTATCTCACTTTGGCCACCACTGCTCGGATPAGTCCAGGCGGCTCAACGACT	856
Oy	1433	TGTTTAGCAAAAAACAACACACGCTGCTGATATTATCATACAGTTTATATAGCTATG	1492
Db	857	TGTTTAGCAAAAAACAACACACGCTGCTGATATTATCATACAGTTTATATAGCTATG	916
Oy	1493	AAGCAGTGAACCGTTTCATTACAGAAATATGATTCAGTGTGCAACGCTGCTAGATGACA	1552
Db	917	AAGCAGTGAACCGTTTCATTACAGAAATATGATTCAGTGTGCAACGCTGCTAGATGACA	976
Oy	1553	CGTCTGGTAAATGA 1566	
Db	977	CGTCTGGTAAATGA 990	

RESULT 7

T38248	
ID	T38248 standard; DNA; 547 bp.
AC	T38248;
DT	28-DEC-1996 (first entry)
DE	Y. pestis F1 antigen cgl gene (including signal sequence).
KE	Plague; vaccine; genetic immunisation; F1 antigen; cgl;
KW	V antigen; ds.
OS	Yersinia pestis strain GB.

Key Location/Qualifiers
 signal_peptide 29..91 /*tag= a
 mat_peptide 92..538 /*tag= b
 WO9628551-A1.
 19-SEP-1996.
 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tlball RM, Williamson ED;
 WPI: 96-433824/43.
 DR P-PSDB; W01043.
 PT Yersinia pestis V antigen and F1 antigen of their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Claim 41; Page 61-62; 96pp; English.
 CC A DNA sequence (T38248) comprises the cafI gene, including the
 CC signal sequence, coding for the F1 antigen (W01043) of Yersinia
 CC pestis. It was obt. by PCR amplification (see also T38257-58)
 CC of Y. pestis DNA. The PCR product was cloned into plasmid pKCMV
 CC and the resulting plasmid (pF1AB) was used to transform E. coli Nova
 CC Blue. Purified plasmid, when administered by i.m. injection,
 CC induced an immunoglobulin response to F1 in BALB/c mice. Live
 CC vaccines comprising gut colonizing organisms transformed with the
 CC cafI gene (see also T38244) can be used to protect a host animal
 CC against plague.
 SQ Sequence 547 BP; 165 A; 120 C; 115 G; 147 T;

Query Match 32.8%; Score 513; DB 1; Length 547;
 Best Local Similarity 100.0%; Pred. No. 1.4e-249;
 Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 68 ATATGAAAAAATCAGTCCGTTATCGCCATTGCAATTATTTGGAATATGCAACTGCTA 127
 DB 27 ATATGAAAAAATCAGTCCGTTATCGCCATTGCAATTATTTGGAATATGCAACTGCTA 86
 0Y 128 ATGGCGGCAATTACTGCAAGCACCACCTGCAAGGCACTCTGTGTGAACAGCCCGCA 187
 DB 87 ATGGCGGCAATTACTGCAAGCACCACCTGCAAGGCACTCTGTGTGAACAGCCCGCA 146
 0Y 188 TCACCTTACATATAAGAAAGCGGCTCAATATGCAATTATTTGGAATATGCAACTGCTA 247
 DB 147 TCACCTTACATATAAGAAAGCGGCTCAATATGCAATTATTTGGAATATGCAACTGCTA 206
 0Y 248 CAGAAATTAATTTGTTAGCTTACTCTTGGGCGCTATTAACAGAACACCTAGACAT 307
 DB 207 CAGAAATTAATTTGTTAGCTTACTCTTGGGCGCTATTAACAGAACACCTAGACAT 266
 0Y 308 CTGTTAACTTACAGATGCGGCGGCTGATCCATGCTTAACTTAACTTAACTTAACTTAACT 367
 DB 267 CTGTTAACTTACAGATGCGGCGGCTGATCCATGCTTAACTTAACTTAACTTAACTTAACT 326
 0Y 368 GAAATTAACCAATTTCACTACAAAAGTATGGCAAGGATCTAGATTTGATATCT 427
 DB 327 GAAATTAACCAATTTCACTACAAAAGTATGGCAAGGATCTAGATTTGATATCT 386
 0Y 428 CTCCTAAGTAAGCGTGAAGACCTTGTGGGATGACCTGCTTGGCTAGAGGCGCCCA 487
 DB 387 CTCCTAAGTAAGCGTGAAGACCTTGTGGGATGACCTGCTTGGCTAGAGGCGCCCA 446
 0Y 488 AGGATTTCTTTTGTGCTAATTTGTTCCAAAGCGGCTTAACTTGAAGAGGATATACA 547
 DB 447 AGGATTTCTTTTGTGCTAATTTGTTCCAAAGCGGCTTAACTTGAAGAGGATATACA 506
 0Y 548 CTGATGCTTAACCGTAACTATTAACCAAG 580
 DB 507 CTGATGCTTAACCGTAACTATTAACCAAG 539

RESULT 8

092819
 ID 092819 standard; DNA; 542 BP.
 AC 092819;
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis cafI (F1) antigen in plasmid pF0RFB.
 KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 KM bubonic plague; pneumonic plague; ds.
 OS Yersinia pestis.
 FH Key
 FT cds
 FT 2..7 Location/Qualifiers
 FT /*tag= a
 FT /note= "first protein encoded by pF0RFB"
 FT 1..6
 FT misc-feature 536..541
 FT /*tag= b
 FT 21..530
 FT /*tag= c
 FT /note= "cafI open reading frame downstream seq."
 FT /*tag= d
 FT /note= "cafI fusion protein"
 PN WO9518231-A1.
 PD 06-JUL-1995.
 PF 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tlball RM, Williamson ED;
 DR WPI: 95-246396/32.
 DR P-PSDB; R76528.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure: Page 19-20; 27pp; English.
 CC The sequence represents the plasmid pF0RFB including the entire
 CC Y. pestis cafI (F1) antigen gene having a 5' tail including a SacI
 CC restriction site, and up to TATAG downstream of the cafI ORF.
 CC The DNA construct can be used to transform human or animal gut
 CC colonizing microorganisms, specifically attenuated Salmonella
 CC typhimurium or Salmonella typhi. The transformed microorganisms
 CC can be used as live/attenuated vaccines which induce immune
 CC responses at mucosal surfaces. The vaccines provide protection
 CC against infection with Y. pestis, and are parenterally and orally
 CC active vaccines offering protection against bubonic and pneumonic
 CC plague.
 SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

Query Match 32.7%; Score 512; DB 1; Length 542;
 Best Local Similarity 100.0%; Pred. No. 4.4e-249;
 Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 68 ATATGAAAAAATCAGTCCGTTATCGCCATTGCAATTATTTGGAATATGCAACTGCTA 127
 DB 19 ATATGAAAAAATCAGTCCGTTATCGCCATTGCAATTATTTGGAATATGCAACTGCTA 78
 0Y 128 ATGGCGGCAATTACTGCAAGCACCACCTGCAAGGCACTCTGTGTGAACAGCCCGCA 187
 DB 79 ATGGCGGCAATTACTGCAAGCACCACCTGCAAGGCACTCTGTGTGAACAGCCCGCA 138
 0Y 188 TCACCTTACATATAAGAAAGCGGCTCAATATGCAATTATTTGGAATATGCAACTGCTA 247
 DB 139 TCACCTTACATATAAGAAAGCGGCTCAATATGCAATTATTTGGAATATGCAACTGCTA 198
 0Y 248 CAGAAATTAATTTGTTAGCTTACTCTTGGGCGCTATTAACAGAACACCTAGACAT 307
 DB 199 CAGAAATTAATTTGTTAGCTTACTCTTGGGCGCTATTAACAGAACACCTAGACAT 258
 0Y 308 CTGTTAACTTACAGATGCGGCGGCTGATCCATGCTTAACTTAACTTAACTTAACTTAACT 367
 DB 259 CTGTTAACTTACAGATGCGGCGGCTGATCCATGCTTAACTTAACTTAACTTAACTTAACT 318
 0Y 368 GAAATTAACCAATTTCACTACAAAAGTATGGCAAGGATCTAGATTTGATATCT 427
 DB 319 GAAATTAACCAATTTCACTACAAAAGTATGGCAAGGATCTAGATTTGATATCT 378

QY 428 CTCCTAAGTAAAGCGTGAGAACCTTGTGGGGGATGACGTGCTTGGCTACGGGACGCC 487
DB 379 CTCCTAAGTAAAGCGTGAGAACCTTGTGGGGGATGACGTGCTTGGCTACGGGACGCC 438
QY 488 AGAATTTCTTGTTCGCTCAATTTGGTCCAAAGCGCGTAACTTGCAGCAGGTAATAACA 547
DB 439 AGAATTTCTTGTTCGCTCAATTTGGTCCAAAGCGCGTAACTTGCAGCAGGTAATAACA 498
QY 548 CTGATGCTGTAAACCGTAAACCGTATCTAACCAA 579
DB 499 CTGATGCTGTAAACCGTAAACCGTATCTAACCAA 530

RESULT 9

VAL1594
ID VAL1594 standard; DNA; 544 Bp.
AC VAL1594:
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nyp1(a)sec544.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 17..529
FT /tag= a
FT /product= "F1 antigen"
PN MO9824912-A2.
PD 11-JUN-1998.
PE 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HEK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR MPI; 98-333331/29.
DR P-PSDB; W59782.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 51; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 544 Bp; 164 A; 120 C; 114 G; 146 T;

Query Match 32.7%; Score 512; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 4.4e-249;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAATCCGTTATCGCATTCGCAATTATTTGGAACTATGCAACTGCTA 127
DB 15 ATATGAAAAAATCAATCCGTTATCGCATTCGCAATTATTTGGAACTATGCAACTGCTA 74
QY 128 ATGGGGAGATTAACTGCAAGCAGCAGCACTGCAAGGCACTCTGTGAACCGCCGCA 187
DB 75 ATGGGGAGATTAACTGCAAGCAGCAGCACTGCAAGGCACTCTGTGAACCGCCGCA 134
QY 188 TCACCTTACATATAAGGAAGCGGCTCCAAATTAATATGCAATGAAACATGATA 247
DB 135 TCACCTTACATATAAGGAAGCGGCTCCAAATTAATATGCAATGAAACATGATA 194
QY 248 CAGAAATTAATGTTGGTACCTTACTCTTGGCGGCTATAAAGCAGAACACATAGCAGAT 307
DB 195 CAGAAATTAATGTTGGTACCTTACTCTTGGCGGCTATAAAGCAGAACACATAGCAGAT 254
QY 308 CTGTAACTTAAAGATGCGCGGGGTATGCCATGTACTTAACATTACTTTCAGATG 367
DB 255 CTGTAACTTAAAGATGCGCGGGGTATGCCATGTACTTAACATTACTTTCAGATG 314
QY 368 GAAATTAACACCAATTCACATAAAGATGATTTGGCAAGGATTCAGAGATTTGATATCT 427
DB 315 GAAATTAACACCAATTCACATAAAGATGATTTGGCAAGGATTCAGAGATTTGATATCT 374

QY 428 CTCCTAAGTAAAGCGTGAGAACCTTGTGGGGGATGACGTGCTTGGCTACGGGACGCC 487
DB 375 CTCCTAAGTAAAGCGTGAGAACCTTGTGGGGGATGACGTGCTTGGCTACGGGACGCC 434
QY 488 AGAATTTCTTGTTCGCTCAATTTGGTCCAAAGCGCGTAACTTGCAGCAGGTAATAACA 547
DB 435 AGAATTTCTTGTTCGCTCAATTTGGTCCAAAGCGCGTAACTTGCAGCAGGTAATAACA 494
QY 548 CTGATGCTGTAAACCGTAAACCGTATCTAACCAA 579
DB 495 CTGATGCTGTAAACCGTAAACCGTATCTAACCAA 526

RESULT 10

VAL1596
ID VAL1596 standard; DNA; 544 Bp.
AC VAL1596:
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nyp1sec510.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 17..532
FT /tag= a
FT /product= "F1 antigen"
PN MO9824912-A2.
PD 11-JUN-1998.
PE 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HEK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR MPI; 98-333331/29.
DR P-PSDB; W59783.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Pages 53-54; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 544 Bp; 166 A; 118 C; 112 G; 148 T;

Query Match 32.7%; Score 512; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 4.4e-249;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAATCCGTTATCGCATTCGCAATTATTTGGAACTATGCAACTGCTA 127
DB 15 ATATGAAAAAATCAATCCGTTATCGCATTCGCAATTATTTGGAACTATGCAACTGCTA 74
QY 128 ATGGGGAGATTAACTGCAAGCAGCAGCACTGCAAGGCACTCTGTGAACCGCCGCA 187
DB 75 ATGGGGAGATTAACTGCAAGCAGCAGCACTGCAAGGCACTCTGTGAACCGCCGCA 134
QY 188 TCACCTTACATATAAGGAAGCGGCTCCAAATTAATATGCAATGAAACATGATA 247
DB 135 TCACCTTACATATAAGGAAGCGGCTCCAAATTAATATGCAATGAAACATGATA 194
QY 248 CAGAAATTAATGTTGGTACCTTACTCTTGGCGGCTATAAAGCAGAACACATAGCAGAT 307
DB 195 CAGAAATTAATGTTGGTACCTTACTCTTGGCGGCTATAAAGCAGAACACATAGCAGAT 254
QY 308 CTGTAACTTAAAGATGCGCGGGGTATGCCATGTACTTAACATTACTTTCAGATG 367
DB 255 CTGTAACTTAAAGATGCGCGGGGTATGCCATGTACTTAACATTACTTTCAGATG 314
QY 368 GAAATTAACACCAATTCACATAAAGATGATTTGGCAAGGATTCAGAGATTTGATATCT 427
DB 315 GAAATTAACACCAATTCACATAAAGATGATTTGGCAAGGATTCAGAGATTTGATATCT 374
QY 428 CTCCTAAGTAAAGCGTGAGAACCTTGTGGGGGATGACGTGCTTGGCTACGGGACGCC 487


```

Db 375 CTCTAAGTAAAGGAGAGACCTGTGGGGGATGAGTGTCTGTGCTACGGGCGACCC 434
    |||||||
OY 488 AGATTTCTTGTGCGTCATGTTTCCAAAGGGGTAACCTGACAGCAATTAATCA 547
    |||||||
Db 435 AGGATTTCTTGTGCGTCATGTTTCCAAAGGGGTAACCTGACAGCAATTAATCA 494
    |||||||
OY 548 CTGATGCTTAACCGTAACCGTATCTAACCA 579
    |||||||
Db 495 CTGATGCTTAACCGTAACCGTATCTAACCA 526
    |||||||

```

RESULT 11

```

V41595
ID V41595 standard; DNA; 510 BP.
AC V41595;
DE 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nYF1(b)sec544.
KW F1 antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
PN MO9824912-A2.
PF 11-JUN-1998.
PR 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-33331/29.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 53; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 510 BP; 153 A; 114 C; 106 G; 137 T;

```

Query Match 32.6%; Score 510; DB 1; Length 510;

Best Local Similarity 100.0%; Pred. No. 4.5e-248; Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 70 ATGAAAAAATCAGTCCGTATGCGCATTCATTTTGGACATTTGCAACTGCTAAT 129
    |||||||
Db 1 ATGAAAAAATCAGTCCGTATGCGCATTCATTTTGGACATTTGCAACTGCTAAT 60
    |||||||
OY 130 GCGGCAATTTAAGTGAAGGCGTCCATTTGCAAGGCACTGTTGTAACGACCGCATC 189
    |||||||
Db 61 GCGGCAATTTAAGTGAAGGCGTCCATTTGCAAGGCACTGTTGTAACGACCGCATC 120
    |||||||
OY 190 ACTTTACATTAAGGAAGGCGTCCATTTGCAAGGCACTGTTGTAACGACCGCATC 249
    |||||||
Db 121 ACTTTACATTAAGGAAGGCGTCCATTTGCAAGGCACTGTTGTAACGACCGCATC 180
    |||||||
OY 250 GAATTAATTTGTTAGGCTTACTCTTGGCGGTATAAAGAGGACACTAGCATCT 309
    |||||||
Db 181 GAATTAATTTGTTAGGCTTACTCTTGGCGGTATAAAGAGGACACTAGCATCT 240
    |||||||
OY 310 GTTAATTTTACAGATCGCGGGGTGATCCATGTAATTAATTTTACTTCTCAGATGA 369
    |||||||
Db 241 GTTAATTTTACAGATCGCGGGGTGATCCATGTAATTAATTTTACTTCTCAGATGA 300
    |||||||
OY 370 AATAACCAACCAATTCACATCAAAAGTGTATGGCAAGATTTAGAGATTTGATCTCT 429
    |||||||
Db 301 AATAACCAACCAATTCACATCAAAAGTGTATGGCAAGATTTAGAGATTTGATCTCT 360
    |||||||
OY 430 CCTAAGGTAACGGTGAAGACCTTGTGGGGGATGACGTCGCTTGGCTAGCGGCGACG 489
    |||||||
Db 361 CCTAAGGTAACGGTGAAGACCTTGTGGGGGATGACGTCGCTTGGCTAGCGGCGACG 420
    |||||||
OY 490 GATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACCTGACAGGATTAATACACT 549
    |||||||
Db 421 GATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACCTGACAGGATTAATACACT 480
    |||||||

```

```

OY 550 GATGCTGAACCGTAACCGTATCTAACCA 579
    |||||||
Db 481 GATGCTGAACCGTAACCGTATCTAACCA 510
    |||||||

```

RESULT 12

```

V41601
ID V41601 standard; DNA; 450 BP.
AC V41601;
DE 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nYF1mat450.
KW F1 antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
PN MO9824912-A2.
PF 11-JUN-1998.
PR 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-33331/29.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 61; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 450 BP; 134 A; 102 C; 96 G; 118 T;

```

Query Match 28.6%; Score 448; DB 1; Length 450;

Best Local Similarity 100.0%; Pred. No. 9.3e-217; Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 132 GCGAATTTAAGTGAAGGCGTCCATTTGCAAGGCACTGTTGTAACGACCGCATC 191
    |||||||
Db 3 GCGAATTTAAGTGAAGGCGTCCATTTGCAAGGCACTGTTGTAACGACCGCATC 62
    |||||||
OY 192 TCTTACATTAAGGAAGGCGTCCATTTGCAAGGCACTGTTGTAACGACCGCATC 251
    |||||||
Db 63 TCTTACATTAAGGAAGGCGTCCATTTGCAAGGCACTGTTGTAACGACCGCATC 122
    |||||||
OY 252 ATTACTTGTGTGCTTACTCTTGGCGGTATAAAGAGGACACTAGCATCTCT 311
    |||||||
Db 123 ATTACTTGTGTGCTTACTCTTGGCGGTATAAAGAGGACACTAGCATCTCT 182
    |||||||
OY 312 TAACTTTACAGATCGCGGGGTGATCCATTTGCAAGGCACTGTTGTAACGACCGCATC 371
    |||||||
Db 183 TAACTTTACAGATCGCGGGGTGATCCATTTGCAAGGCACTGTTGTAACGACCGCATC 242
    |||||||
OY 372 TAACCAACCAATTCACATCAAAAGTGTATGGCAAGATTTAGAGATTTGATCTCTCC 431
    |||||||
Db 243 TAACCAACCAATTCACATCAAAAGTGTATGGCAAGATTTAGAGATTTGATCTCTCC 302
    |||||||
OY 432 TAAAGTAACCGTGAAGACCTTGTGGGGGATGACGTCGCTTGGCTAGCGGCGACG 491
    |||||||
Db 303 TAAAGTAACCGTGAAGACCTTGTGGGGGATGACGTCGCTTGGCTAGCGGCGACG 362
    |||||||
OY 492 TTTGTTTGTGTCATTTGTTCCAAAGCGGTAACCTGACAGGATTAATACACTGA 551
    |||||||
Db 363 TTTGTTTGTGTCATTTGTTCCAAAGCGGTAACCTGACAGGATTAATACACTGA 422
    |||||||
OY 552 TGCTGTAAACCGTAACCGTATCTAACCA 579
    |||||||
Db 423 TGCTGTAAACCGTAACCGTATCTAACCA 450
    |||||||

```

RESULT 13

```

ID V41600
V41600 standard; DNA; 474 BP.
AC V41600;

```

26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nyP1mat474.
 KW F1 antigen; plasmid; vaccine; plaque; ds.
 OS Yersinia pestis.
 FT Key Location/Qualifiers
 FT CDS 7..459
 FT /tag= a
 FT /product= "F1 antigen"
 WO9824912-A2.
 PN 11-JUN-1998.
 PD 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-333331/29.
 DR P-PSDB: W59787.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Pages 59-60; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 CC Sequence 474 BP; 143 A; 106 C; 99 G; 126 T;

Query Match 28.6%; Score 448; DB 1; Length 474;
 Best Local Similarity 100.0%; Pred. No. 9.3e-217; Mismatches 0; Indels 0; Gaps 0;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 GGCAGATTTAACTGCAAGCACCACGCAAGGCACTCTTTGTAACCAAGCCGCAATCAC 191
 DB 9 GGCAGATTTAACTGCAAGCACCACGCAAGGCACTCTTTGTAACCAAGCCGCAATCAC 68
 QY 192 TCTTAATATTAAGAGAGCGCCCTCAATTAATATGAGCAATGGAACATCGATACAGA 251
 DB 69 TCTTAATATTAAGAGAGCGCCCTCAATTAATATGAGCAATGGAACATCGATACAGA 128
 QY 252 ATTACTTGTGTGAGCCTTACTCTTGGCGGCTATTAAGAGCAAGCACTGACATCTGT 311
 DB 129 ATTACTTGTGTGAGCCTTACTCTTGGCGGCTATTAAGAGCAAGCACTGACATCTGT 188
 QY 312 TAACTTTACAGATGCGCGGGGTGATCCCATGTAATTAATTAATCTCTCAGAGATGAAA 371
 DB 189 TAACTTTACAGATGCGCGGGGTGATCCCATGTAATTAATTAATCTCTCAGAGATGAAA 248
 QY 372 TAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 431
 DB 249 TAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 308
 QY 432 TAAAGTAAGAGGATGAGACCTTGTGGGGATGAGCTGCTTGGCTACAGGAGCCAGAGA 491
 DB 309 TAAAGTAAGAGGATGAGACCTTGTGGGGATGAGCTGCTTGGCTACAGGAGCCAGAGA 368
 QY 492 TTTCTTTGTTGCTCAATTTGGTTCCAAAGCGGTTAACTTGACAGAGGTAATATACACTGA 551
 DB 369 TTTCTTTGTTGCTCAATTTGGTTCCAAAGCGGTTAACTTGACAGAGGTAATATACACTGA 428
 QY 552 TGCTGTACCGGTACCGGTATCTAACCA 579
 DB 429 TGCTGTACCGGTACCGGTATCTAACCA 456

RESULT 14
 ID V41609 standard; DNA; 447 BP.
 AC V41609;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nyP1mat447.
 KW F1 antigen; plasmid; vaccine; plaque; ds.
 OS Yersinia pestis.
 FT Key Location/Qualifiers
 FT CDS 1..576
 FT /tag= a
 FT /product= "F1 antigen"

1..447
 FT CDS /tag= a
 FT /product= "F1 antigen"
 WO9824912-A2.
 PN 11-JUN-1998.
 PD 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-333331/29.
 DR P-PSDB: W59788.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Page 64; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 CC Sequence 447 BP; 133 A; 102 C; 95 G; 117 T;

Query Match 28.5%; Score 447; DB 1; Length 447;
 Best Local Similarity 100.0%; Pred. No. 3e-216; Mismatches 0; Indels 0; Gaps 0;
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 GCAGATTTAACTGCAAGCACCACGCAAGGCACTCTTTGTAACCAAGCCGCAATCAC 192
 DB 1 GCAGATTTAACTGCAAGCACCACGCAAGGCACTCTTTGTAACCAAGCCGCAATCAC 60
 QY 193 CTTCATATTAAGAGAGCGCCCTCAATTAATATGAGCAATGGAACATCGATACAGA 252
 DB 61 CTTCATATTAAGAGAGCGCCCTCAATTAATATGAGCAATGGAACATCGATACAGA 120
 QY 253 TTAATTTGTTGCTCAATTTGGTTCCAAAGCGGTTAACTTGACAGAGGTAATATACACTGA 312
 DB 121 TTAATTTGTTGCTCAATTTGGTTCCAAAGCGGTTAACTTGACAGAGGTAATATACACTGA 180
 QY 313 AACCTTAAGATGCGCGGGGTGATCCCATGTAATTAATTAATTAATTAATTAATTAAT 372
 DB 181 AACCTTAAGATGCGCGGGGTGATCCCATGTAATTAATTAATTAATTAATTAATTAAT 240
 QY 373 AACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 432
 DB 241 AACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
 QY 433 AAGGTAAAGGATGAGACCTTGTGGGGATGAGCTGCTTGGCTACAGGAGCCAGAGAT 492
 DB 301 AAGGTAAAGGATGAGACCTTGTGGGGATGAGCTGCTTGGCTACAGGAGCCAGAGAT 360
 QY 493 TTTCTTTGTTGCTCAATTTGGTTCCAAAGCGGTTAACTTGACAGAGGTAATATACACTGA 552
 DB 361 TTTCTTTGTTGCTCAATTTGGTTCCAAAGCGGTTAACTTGACAGAGGTAATATACACTGA 420
 QY 553 GCTGTACCGGTACCGGTATCTAACCA 579
 DB 421 GCTGTACCGGTACCGGTATCTAACCA 447

RESULT 15
 ID V41598 standard; DNA; 576 BP.
 AC V41598;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nyP1anc576.
 KW F1 antigen; plasmid; vaccine; plaque; ds.
 OS Yersinia pestis.
 FT Key Location/Qualifiers
 FT CDS 1..576
 FT /tag= a
 FT /product= "F1 antigen"
 WO9824912-A2.
 PN 11-JUN-1998.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 15:36:43 ; Search time 46.98 Seconds
(without alignments)
4585.194 Million cell updates/sec

Title: US-08-699-716a-1

Perfect score: 1566
Sequence: 1 ATGGGCGCATCATCATCATCA.....ATGCACGCTGCTGAATGCA 1566

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 243080 seqs, 68777915 residues

Word size: 0

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database:

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/1na/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/1na/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/1na/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/1na/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/1na/6_COMB.seq:*
6: /cgn2_6/ptodata/1/1na/6CTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/1na/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	927	59.2	1014	4 US-08-913-477-1	Sequence 1, Appl
2	927	59.2	1462	4 US-08-913-477-16	Sequence 16, Appl
3	927	59.2	1530	4 US-08-913-477-22	Sequence 22, Appl
4	923	58.9	1014	4 US-08-913-477-3	Sequence 3, Appl
5	411	26.2	547	4 US-08-913-477-20	Sequence 20, Appl
6	398	25.4	541	4 US-08-913-477-10	Sequence 10, Appl
7	68	4.3	375	5 US-08-870-370-8	Sequence 8, Appl
8	68	4.3	375	5 US-08-870-370-9	Sequence 9, Appl
9	68	4.3	375	5 US-08-870-370-8	Sequence 8, Appl
10	53	3.4	1402	2 US-08-480-604A-25	Sequence 25, Appl
11	53	3.4	1402	4 US-08-405-486A-25	Sequence 25, Appl
12	35	2.2	47	4 US-08-600-783-15	Sequence 25, Appl
13	30	1.9	8920	3 US-08-446-855A-1	Sequence 15, Appl
14	29	1.9	4084	5 US-08-866-340-1	Sequence 1, Appl
15	28	1.8	39	3 US-08-432-871C-44	Sequence 44, Appl
16	28	1.8	88	6 PCT-US95-10973A-36	Sequence 36, Appl
17	28	1.8	287	4 US-08-347-563A-11	Sequence 11, Appl
18	28	1.8	287	4 US-08-292-345B-11	Sequence 11, Appl
19	28	1.8	287	4 US-08-485-942A-11	Sequence 11, Appl
20	28	1.8	732	6 PCT-US96-05611A-29	Sequence 29, Appl
21	28	1.8	2136	5 US-08-946-475-8	Sequence 8, Appl
22	28	1.8	6693	4 US-08-147-777-2	Sequence 2, Appl
23	28	1.8	6693	5 US-08-452-872-2	Sequence 2, Appl
24	28	1.8	6693	6 PCT-US93-03985-2	Sequence 2, Appl
25	28	1.8	10747	4 US-08-147-777-1	Sequence 1, Appl
26	28	1.8	10747	5 US-08-452-872-1	Sequence 1, Appl

27	28	1.8	10747	6 PCT-US93-03985-1	Sequence 1, Appl
28	28	1.8	24979	4 US-08-147-777-3	Sequence 3, Appl
29	28	1.8	24979	5 US-08-452-872-3	Sequence 3, Appl
30	28	1.8	24979	6 PCT-US93-03985-3	Sequence 3, Appl
31	25	1.6	225	4 US-08-913-477-9	Sequence 9, Appl
32	24	1.5	1667	2 US-08-485-284A-1	Sequence 1, Appl
33	23	1.5	32	4 US-08-913-477-7	Sequence 7, Appl
34	23	1.5	797	3 US-08-752-132-1	Sequence 1, Appl
35	23	1.5	5434	4 US-08-841-349-1	Sequence 1, Appl
36	23	1.5	7635	2 US-08-455-543A-1	Sequence 1, Appl
37	23	1.5	7635	3 US-08-455-543A-23	Sequence 23, Appl
38	23	1.5	7635	3 US-08-193-078B-1	Sequence 23, Appl
39	23	1.5	7635	3 US-08-193-078B-29	Sequence 29, Appl
40	23	1.5	7635	3 US-08-223-305C-1	Sequence 1, Appl
41	23	1.5	7635	3 US-08-223-305C-23	Sequence 23, Appl
42	23	1.5	7635	3 US-08-149-097D-1	Sequence 1, Appl
43	22	1.4	2625	3 US-08-759-945-1	Sequence 1, Appl
44	22	1.4	7125	3 US-07-745-206A-1	Sequence 1, Appl
45	22	1.4	7125	3 US-08-311-363-1	Sequence 1, Appl
46	21	1.3	21	4 US-08-863-639A-44	Sequence 44, Appl
47	21	1.3	21	4 US-08-863-639A-45	Sequence 45, Appl
48	21	1.3	21	4 US-08-863-639A-49	Sequence 49, Appl
49	21	1.3	21	4 US-08-863-639A-65	Sequence 65, Appl
50	21	1.3	21	4 US-08-863-639A-82	Sequence 82, Appl
51	21	1.3	21	4 US-08-863-639A-86	Sequence 86, Appl
52	21	1.3	47	3 US-08-762-433-5	Sequence 5, Appl
53	21	1.3	47	3 US-09-001-219-5	Sequence 5, Appl
54	21	1.3	1475	2 US-08-097-938-1	Sequence 1, Appl
55	21	1.3	1475	2 US-08-476-000-1	Sequence 1, Appl
56	21	1.3	1475	2 US-08-472-840-1	Sequence 1, Appl
57	21	1.3	1475	3 US-08-472-840-1	Sequence 1, Appl
58	21	1.3	1475	5 US-08-476-976-1	Sequence 1, Appl
59	21	1.3	1497	6 PCT-US96-06511-5	Sequence 5, Appl
60	21	1.3	1497	6 US-08-474-410-1	Sequence 5, Appl
61	21	1.3	2732	2 US-08-476-000-50	Sequence 60, Appl
62	21	1.3	2732	2 US-08-472-840-50	Sequence 60, Appl
63	21	1.3	2732	3 US-08-476-976-60	Sequence 60, Appl
64	21	1.3	2732	5 US-08-474-410-60	Sequence 60, Appl
65	20	1.3	28	4 US-08-913-477-19	Sequence 19, Appl
66	20	1.3	40	1 US-07-743-245-1	Sequence 1, Appl
67	20	1.3	1297	7 5187077-23	Patent No. 5187077
68	20	1.3	1297	7 5427925-21	Patent No. 5427925
69	20	1.3	2469	1 US-07-997-133-2	Sequence 2, Appl
70	20	1.3	2469	1 US-08-459-286-1	Sequence 1, Appl
71	20	1.3	2469	6 US-07-997-133-2	Sequence 2, Appl
72	20	1.3	2662	3 US-08-451-822A-14	Sequence 14, Appl
73	20	1.3	2733	2 US-08-371-001-14	Sequence 14, Appl
74	20	1.3	2733	6 PCT-US96-00331-14	Sequence 14, Appl
75	20	1.3	5194	3 US-08-642-846-1	Sequence 1, Appl
76	20	1.2	30	2 US-08-737-757-7	Sequence 7, Appl
77	19	1.2	30	2 US-08-737-757-8	Sequence 8, Appl
78	19	1.2	33	2 US-08-482-882-29	Sequence 29, Appl
79	19	1.2	33	2 US-08-483-389-29	Sequence 29, Appl
80	19	1.2	33	3 US-08-487-113D-29	Sequence 29, Appl
81	19	1.2	33	3 US-08-473-503-29	Sequence 29, Appl
82	19	1.2	33	3 US-08-483-922-29	Sequence 29, Appl
83	19	1.2	33	4 US-08-720-400A-29	Sequence 29, Appl
84	19	1.2	33	4 US-08-714-017-29	Sequence 29, Appl
85	19	1.2	40	2 US-08-482-882-30	Sequence 30, Appl
86	19	1.2	40	2 US-08-483-389-30	Sequence 30, Appl
87	19	1.2	40	3 US-08-487-113D-30	Sequence 30, Appl
88	19	1.2	40	3 US-08-473-503-30	Sequence 30, Appl
89	19	1.2	40	3 US-08-483-922-30	Sequence 30, Appl
90	19	1.2	40	4 US-08-483-922-30	Sequence 30, Appl
91	19	1.2	40	4 US-08-720-400A-30	Sequence 30, Appl
92	19	1.2	41	2 US-08-714-017-30	Sequence 30, Appl
93	19	1.2	45	5 US-08-525-654A-139	Sequence 139, App
94	19	1.2	48	4 US-08-481-435-19	Sequence 19, Appl
95	19	1.2	48	4 US-09-061-337-19	Sequence 19, Appl
96	19	1.2	49	4 US-09-122-139-19	Sequence 19, Appl
97	19	1.2	49	5 US-09-340-991-19	Sequence 19, Appl
98	19	1.2	51	4 US-09-061-337-13	Sequence 13, Appl
99	19	1.2	51	4 US-09-122-129-13	Sequence 13, Appl
				5 US-09-340-991-13	Sequence 13, Appl

100	19	1.2	53	1	US-08-049-264C-12	Sequence 12, Appl	173	18	1.1	3661	1	US-08-105-483-221	Sequence 221, App
101	19	1.2	53	1	US-08-476-562-12	Sequence 12, Appl	174	18	1.1	3661	1	US-08-303-124-7	Sequence 7, Appl1
102	19	1.2	53	1	US-08-479-723A-12	Sequence 12, Appl	175	18	1.1	3661	1	US-08-204-729-7	Sequence 7, Appl1
103	19	1.2	53	6	PCT-US94-04310-12	Sequence 12, Appl	176	18	1.1	3661	2	US-08-709-209-221	Sequence 221, App
104	19	1.2	57	4	US-08-662-227-39	Sequence 39, Appl	177	18	1.1	3661	2	US-08-458-101-221	Sequence 221, App
105	19	1.2	96	5	US-08-737-336-2	Sequence 23, Appl	178	18	1.1	3661	2	US-08-480-697B-7	Sequence 7, Appl1
106	19	1.2	105	4	US-08-687-865A-23	Sequence 23, Appl	179	18	1.1	4453	3	US-08-770-761A-4	Sequence 6, Appl1
107	19	1.2	129	4	US-08-679-865-38	Sequence 38, Appl	180	18	1.1	4453	2	US-08-770-761A-6	Sequence 6, Appl1
108	19	1.2	129	4	US-08-680-876-38	Sequence 38, Appl	181	18	1.1	6387	2	US-07-721-775A-1	Sequence 1, Appl1
109	19	1.2	153	4	US-08-679-865-41	Sequence 41, Appl	182	18	1.1	6387	1	US-08-339-658-1	Sequence 1, Appl1
110	19	1.2	153	4	US-08-680-876-41	Sequence 41, Appl	183	17	1.1	27	7	5463174-1	Patent No. 5463174
111	19	1.2	234	4	US-08-687-865A-20	Sequence 20, Appl	184	17	1.1	27	7	5463174-2	Patent No. 5463174
112	19	1.2	466	1	US-08-457-245-20	Sequence 20, Appl	185	17	1.1	41	3	US-08-632-877C-43	Sequence 43, Appl
113	19	1.2	927	1	US-08-499-568-1	Sequence 1, Appl1	186	17	1.1	303	5	US-08-621-018B-5	Sequence 5, Appl1
114	19	1.2	927	2	US-08-793-958A-1	Sequence 1, Appl1	187	17	1.1	1338	4	US-08-993-228-5	Sequence 25, Appl1
115	19	1.2	1056	4	US-08-701-191A-5	Sequence 5, Appl1	188	17	1.1	1183	1	US-08-188-582-25	Sequence 25, Appl
116	19	1.2	1240	4	US-09-061-337-9	Sequence 9, Appl1	189	17	1.1	1183	1	US-08-646-715-25	Sequence 25, Appl
117	19	1.2	1240	4	US-09-122-129-9	Sequence 9, Appl1	190	17	1.1	1282	1	US-08-211-942-16	Sequence 16, Appl
118	19	1.2	1240	5	US-09-340-591-9	Sequence 9, Appl1	191	17	1.1	1328	1	US-08-290-448A-58	Sequence 58, Appl
119	19	1.2	1240	5	US-08-737-336-5	Sequence 5, Appl1	192	17	1.1	1328	2	US-08-175-069A-58	Sequence 58, Appl
120	19	1.2	1740	4	US-08-335-760-3	Sequence 3, Appl1	193	17	1.1	1328	2	US-08-290-448A-58	Sequence 58, Appl
121	19	1.2	1896	2	US-08-605-541B-11	Sequence 11, Appl	194	17	1.1	1349	2	US-08-290-448A-73	Sequence 73, Appl
122	19	1.2	2081	4	US-09-096-982-7	Sequence 7, Appl	195	17	1.1	1349	2	US-08-290-448A-73	Sequence 73, Appl
123	19	1.2	2081	4	US-08-653-650A-7	Sequence 7, Appl1	196	17	1.1	1349	2	US-08-175-069A-73	Sequence 73, Appl
124	19	1.2	3090	6	PCT-US93-06251-7	Sequence 5, Appl1	197	17	1.1	1349	2	US-08-221-235-6	Sequence 6, Appl1
125	19	1.2	3255	2	US-08-717-515-5	Sequence 5, Appl1	198	17	1.1	1349	2	US-08-903-801-2	Sequence 2, Appl1
126	19	1.2	3704	4	US-09-014-969-20	Sequence 20, Appl	199	17	1.1	2120	5	US-09-221-235-4	Sequence 4, Appl1
127	19	1.2	3831	2	US-08-717-515-7	Sequence 7, Appl	200	17	1.1	2188	1	US-07-865-662F-10	Sequence 10, Appl
128	19	1.2	6934	4	US-08-015-973-2	Sequence 2, Appl1	201	17	1.1	2295	5	US-08-375-300-3	Sequence 3, Appl1
129	19	1.2	6934	4	US-08-448-164-2	Sequence 2, Appl1	202	17	1.1	2295	5	US-09-177-431-3	Sequence 3, Appl1
130	19	1.2	7286	6	PCT-US95-11684-1	Sequence 1, Appl1	203	17	1.1	2295	5	PCT-US95-16930-3	Sequence 3, Appl1
131	19	1.2	18	2	US-08-470-837-21	Sequence 21, Appl	204	17	1.1	3074	6	PCT-US93-11725-1	Sequence 6, Appl1
132	18	1.1	18	5	US-09-031-897-10	Sequence 10, Appl	205	17	1.1	3707	1	US-08-118-101A-1	Sequence 1, Appl1
133	18	1.1	35	1	US-08-049-562C-10	Sequence 10, Appl	206	17	1.1	4080	5	US-08-375-300-1	Sequence 5, Appl1
134	18	1.1	35	1	US-08-476-562-10	Sequence 10, Appl	207	17	1.1	4080	5	US-09-177-431-1	Sequence 1, Appl1
135	18	1.1	35	6	US-08-479-723A-10	Sequence 10, Appl	208	17	1.1	4080	6	PCT-US95-16930-1	Sequence 1, Appl1
136	18	1.1	35	6	PCT-US94-04310-10	Sequence 10, Appl	209	17	1.1	5183	3	US-08-459-568-3	Sequence 3, Appl1
137	18	1.1	40	3	US-08-713-815A-6	Sequence 6, Appl1	210	17	1.1	5183	3	US-08-399-411-3	Sequence 3, Appl1
138	18	1.1	43	5	US-09-136-421-1	Sequence 1, Appl1	211	17	1.1	5718	5	US-08-714-918-48	Sequence 48, Appl
139	18	1.1	45	5	US-08-481-435-40	Sequence 40, Appl	212	17	1.1	5668	5	US-08-516-859A-3	Sequence 3, Appl1
140	18	1.1	49	2	US-08-052-157-1	Sequence 1, Appl1	213	17	1.1	9468	2	US-08-325-547-10	Sequence 10, Appl
141	18	1.1	51	1	US-08-442-134A-4	Sequence 4, Appl1	214	17	1.1	9472	2	US-08-325-547-9	Sequence 9, Appl1
142	18	1.1	51	1	US-08-444-581B-4	Sequence 4, Appl1	215	17	1.1	13875	3	US-08-734-344-1	Sequence 1, Appl1
143	18	1.1	51	2	US-08-446-088A-4	Sequence 4, Appl1	216	16	1.0	21	5	US-08-851-350-31	Sequence 31, Appl
144	18	1.1	51	3	US-08-642-541-7	Sequence 7, Appl1	217	16	1.0	21	5	US-08-851-350-32	Sequence 32, Appl
145	18	1.1	53	4	US-08-687-865A-10	Sequence 10, Appl	218	16	1.0	24	4	US-08-600-783-16	Sequence 16, Appl
146	18	1.1	55	4	US-08-687-865A-11	Sequence 11, Appl	219	16	1.0	29	1	US-08-276-852-140	Sequence 140, App
147	18	1.1	55	4	US-08-484-993B-39	Sequence 39, Appl	220	16	1.0	29	1	US-08-276-852-140	Sequence 140, App
148	18	1.1	58	4	US-08-484-158B-39	Sequence 39, Appl	221	16	1.0	29	2	US-08-899-575-139	Sequence 139, App
149	18	1.1	58	4	US-08-484-596A-39	Sequence 39, Appl	222	16	1.0	29	2	US-08-899-575-140	Sequence 140, App
150	18	1.1	58	4	US-08-480-150A-39	Sequence 39, Appl	223	16	1.0	29	2	US-08-899-575-139	Sequence 139, App
151	18	1.1	58	5	US-08-458-731-39	Sequence 39, Appl	224	16	1.0	29	2	US-08-899-575-140	Sequence 140, App
152	18	1.1	58	5	US-08-149-223A-39	Sequence 39, Appl	225	16	1.0	29	6	PCT-US95-08743-139	Sequence 139, App
153	18	1.1	688	2	US-08-139-937-9	Sequence 9, Appl1	226	16	1.0	29	6	PCT-US95-08743-140	Sequence 140, App
154	18	1.1	688	6	PCT-US93-11310-9	Sequence 9, Appl1	227	16	1.0	39	5	US-08-792-832A-8	Sequence 8, Appl1
155	18	1.1	1293	3	US-08-722-184-7	Sequence 7, Appl1	228	16	1.0	42	5	US-09-136-421-2	Sequence 2, Appl1
156	18	1.1	2109	3	US-08-617-801A-5	Sequence 5, Appl1	229	16	1.0	42	5	US-09-136-421-7	Sequence 7, Appl1
157	18	1.1	2220	3	US-08-617-801A-1	Sequence 1, Appl1	230	16	1.0	141	5	US-09-003-708A-2	Sequence 2, Appl1
158	18	1.1	2722	4	US-08-500-857A-7	Sequence 7, Appl1	231	16	1.0	220	1	US-08-276-852-11	Sequence 11, Appl
159	18	1.1	2809	1	US-08-484-105-3	Sequence 3, Appl1	232	16	1.0	220	1	US-08-276-852-12	Sequence 12, Appl
160	18	1.1	2809	1	US-08-484-106-3	Sequence 3, Appl1	233	16	1.0	220	1	US-08-899-575-11	Sequence 11, Appl
161	18	1.1	3635	3	US-08-553-436A-5	Sequence 5, Appl1	234	16	1.0	220	2	US-08-899-575-12	Sequence 12, Appl
162	18	1.1	3635	3	US-08-220-151-72	Sequence 72, Appl	235	16	1.0	220	2	US-08-899-575-11	Sequence 11, Appl
163	18	1.1	3659	1	US-08-303-124-12	Sequence 12, Appl	236	16	1.0	220	2	US-08-899-575-12	Sequence 12, Appl
164	18	1.1	3659	1	US-08-413-118-72	Sequence 72, Appl	237	16	1.0	220	6	PCT-US95-08743-11	Sequence 11, Appl
165	18	1.1	3659	2	US-08-324-657-48	Sequence 48, Appl	238	16	1.0	220	6	PCT-US95-08743-12	Sequence 12, Appl
166	18	1.1	3659	3	US-08-184-009-77	Sequence 77, Appl	239	16	1.0	315	4	US-08-487-031-21	Sequence 21, Appl
167	18	1.1	3659	3	US-08-417-210A-67	Sequence 67, Appl	240	16	1.0	534	2	US-08-700-575-1	Sequence 1, Appl1
168	18	1.1	3659	3	US-08-480-697B-12	Sequence 12, Appl	241	16	1.0	534	2	US-08-487-031-22	Sequence 22, Appl
169	18	1.1	3659	4	US-08-458-356-77	Sequence 77, Appl	242	16	1.0	604	4	US-08-487-031-19	Sequence 19, Appl
170	18	1.1	3659	5	US-08-473-446-72	Sequence 72, Appl	243	16	1.0	604	4	US-08-836-197-1	Sequence 1, Appl1
171	18	1.1	3660	2	US-08-475-063-25	Sequence 25, Appl	244	16	1.0	677	2	US-08-644-664B-3	Sequence 3, Appl1
172	18	1.1	3660	2	US-08-207-792-25	Sequence 25, Appl	245	16	1.0	677	4	US-08-761-277-3	Sequence 3, Appl1

C 246	16	1.0	909	4	US-08-487-031-18	Sequence 19, Appl	C 319	15	1.0	21	5	US-08-840-316-28	Sequence 28, Appl
C 247	16	1.0	1016	1	US-08-109-391A-3	Sequence 3, Appl1	C 320	15	1.0	21	6	PCT-US93-08848A-28	Sequence 28, Appl
C 248	16	1.0	1016	1	US-08-459-019A-3	Sequence 3, Appl1	C 321	15	1.0	21	6	PCT-US93-08848A-28	Sequence 28, Appl
C 249	16	1.0	1016	1	US-08-460-428A-3	Sequence 3, Appl1	C 322	15	1.0	21	6	PCT-US93-08848A-28	Sequence 28, Appl
C 250	16	1.0	1033	3	US-08-471-717-3	Sequence 3, Appl1	C 323	15	1.0	38	2	US-08-255-272-31	Sequence 31, Appl
C 251	16	1.0	1158	2	US-08-891-254-4	Sequence 4, Appl1	C 324	15	1.0	38	2	US-08-317-832A-32	Sequence 32, Appl
C 252	16	1.0	1158	3	US-08-819-539-4	Sequence 4, Appl1	C 325	15	1.0	56	5	US-08-792-832A-17	Sequence 17, Appl
C 253	16	1.0	1158	6	PCT-US93-06243-4	Sequence 4, Appl1	C 326	15	1.0	57	5	US-08-634-060-30	Sequence 30, Appl
C 254	16	1.0	1158	6	PCT-US96-08819-4	Sequence 4, Appl1	C 327	15	1.0	81	1	US-07-939-501A-29	Sequence 29, Appl
C 255	16	1.0	1165	5	US-08-759-463-1	Sequence 1, Appl1	C 328	15	1.0	109	1	US-07-939-501A-27	Sequence 27, Appl
C 256	16	1.0	1185	1	US-07-876-380-8	Sequence 8, Appl1	C 329	15	1.0	176	4	US-08-687-080-82	Sequence 82, Appl
C 257	16	1.0	1185	1	US-08-049-783-5	Sequence 5, Appl1	C 330	15	1.0	247	2	US-08-691-641-4	Sequence 4, Appl1
C 258	16	1.0	1185	2	US-08-316-301A-9	Sequence 9, Appl1	C 331	15	1.0	305	1	US-08-207-904-13	Sequence 13, Appl
C 259	16	1.0	1185	2	US-08-904-278-9	Sequence 9, Appl1	C 332	15	1.0	342	2	US-08-486-013-56	Sequence 56, Appl
C 260	16	1.0	1189	6	PCT-US92-03624-9	Sequence 9, Appl1	C 333	15	1.0	342	3	US-08-482-279-56	Sequence 56, Appl
C 261	16	1.0	1189	1	US-07-781-034-4	Sequence 4, Appl1	C 334	15	1.0	342	3	US-08-342-268-56	Sequence 56, Appl
C 262	16	1.0	1189	6	PCT-US92-08328-4	Sequence 4, Appl1	C 335	15	1.0	342	3	US-08-342-268-56	Sequence 56, Appl
C 263	16	1.0	1203	3	US-08-739-945-2	Sequence 2, Appl1	C 336	15	1.0	352	5	US-09-015-568-56	Sequence 56, Appl
C 264	16	1.0	1203	3	US-08-200-724A-4	Sequence 4, Appl1	C 337	15	1.0	355	2	US-08-486-013-43	Sequence 43, Appl
C 265	16	1.0	1227	3	US-08-683-007A-1	Sequence 1, Appl1	C 338	15	1.0	355	3	US-08-482-279-43	Sequence 43, Appl
C 266	16	1.0	1287	6	PCT-US93-06243-3	Sequence 3, Appl1	C 339	15	1.0	355	3	US-08-342-268-43	Sequence 43, Appl
C 267	16	1.0	1288	3	US-08-200-724A-3	Sequence 3, Appl1	C 340	15	1.0	355	5	US-09-015-568-43	Sequence 43, Appl
C 268	16	1.0	1288	4	US-08-030-270A-4	Sequence 4, Appl1	C 341	15	1.0	406	5	US-08-928-799A-1	Sequence 1, Appl1
C 269	16	1.0	1320	1	US-08-250-448A-75	Sequence 75, Appl	C 342	15	1.0	419	2	US-08-486-013-62	Sequence 62, Appl
C 270	16	1.0	1320	2	US-08-290-448A-75	Sequence 75, Appl	C 343	15	1.0	419	3	US-08-482-279-62	Sequence 62, Appl
C 271	16	1.0	1320	2	US-08-175-069A-75	Sequence 75, Appl	C 344	15	1.0	419	5	US-09-015-568-62	Sequence 62, Appl
C 272	16	1.0	1400	3	US-08-001-078A-2	Sequence 2, Appl1	C 345	15	1.0	422	2	US-08-486-013-41	Sequence 41, Appl
C 273	16	1.0	1400	4	US-08-463-218-2	Sequence 2, Appl1	C 346	15	1.0	422	2	US-08-486-013-41	Sequence 41, Appl
C 274	16	1.0	1400	6	PCT-US94-00253-2	Sequence 2, Appl1	C 347	15	1.0	422	3	US-08-482-279-41	Sequence 41, Appl
C 275	16	1.0	1697	1	US-08-181-271A-104	Sequence 104, App	C 348	15	1.0	422	3	US-08-482-279-45	Sequence 45, Appl
C 276	16	1.0	1697	1	US-08-449-315-104	Sequence 104, App	C 349	15	1.0	422	3	US-08-482-279-45	Sequence 45, Appl
C 277	16	1.0	1697	1	US-08-444-803-104	Sequence 104, App	C 350	15	1.0	422	3	US-08-342-268-45	Sequence 45, Appl
C 278	16	1.0	1697	1	US-08-449-043-104	Sequence 104, App	C 351	15	1.0	422	3	US-08-342-268-45	Sequence 45, Appl
C 279	16	1.0	1697	2	US-08-456-262-104	Sequence 104, App	C 352	15	1.0	422	5	US-09-015-568-45	Sequence 45, Appl
C 280	16	1.0	1697	2	US-08-455-416-104	Sequence 104, App	C 353	15	1.0	422	5	US-09-015-568-45	Sequence 45, Appl
C 281	16	1.0	1697	2	US-08-455-416-104	Sequence 104, App	C 354	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 282	16	1.0	1697	2	US-08-455-416-104	Sequence 104, App	C 355	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 283	16	1.0	1697	2	US-08-455-416-104	Sequence 104, App	C 356	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 284	16	1.0	1697	3	US-08-456-262-104	Sequence 104, App	C 357	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 285	16	1.0	1697	3	US-08-456-262-104	Sequence 104, App	C 358	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 286	16	1.0	1697	3	US-08-455-416-104	Sequence 104, App	C 359	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 287	16	1.0	1697	3	US-08-455-416-104	Sequence 104, App	C 360	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 288	16	1.0	1697	4	US-08-971-217-104	Sequence 104, App	C 361	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 289	16	1.0	1697	4	US-08-185-282-104	Sequence 104, App	C 362	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 290	16	1.0	2091	1	US-08-286-856C-1	Sequence 1, Appl1	C 363	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 291	16	1.0	2091	1	US-08-472-831-1	Sequence 1, Appl1	C 364	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 292	16	1.0	2133	3	US-08-577-922-31	Sequence 31, Appl	C 365	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 293	16	1.0	2133	3	US-08-577-922-31	Sequence 31, Appl	C 366	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 294	16	1.0	2467	5	US-08-955-937A-1	Sequence 1, Appl1	C 367	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 295	16	1.0	2510	5	US-08-787-739-58	Sequence 58, Appl	C 368	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 296	16	1.0	2510	4	US-08-888-982A-42	Sequence 42, Appl	C 369	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 297	16	1.0	2790	1	US-08-018-129-4	Sequence 4, Appl1	C 370	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 298	16	1.0	2814	4	US-08-448-250-4	Sequence 4, Appl1	C 371	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 299	16	1.0	3141	4	US-08-743-637B-186	Sequence 186, App	C 372	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 300	16	1.0	3141	4	US-08-956-242-1	Sequence 1, Appl1	C 373	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 301	16	1.0	3141	4	US-08-956-242-1	Sequence 1, Appl1	C 374	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 302	16	1.0	3171	4	US-08-868-786-5	Sequence 5, Appl1	C 375	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 303	16	1.0	3383	1	US-07-707-367-1	Sequence 1, Appl1	C 376	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 304	16	1.0	3441	4	US-08-525-864A-1	Sequence 1, Appl1	C 377	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 305	16	1.0	3532	4	US-08-767-739-9	Sequence 9, Appl1	C 378	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 306	16	1.0	5243	3	US-09-039-555B-18	Sequence 18, Appl	C 379	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 307	16	1.0	6008	3	US-08-414-335-2	Sequence 2, Appl1	C 380	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 308	16	1.0	6008	3	US-07-789-915A-5	Sequence 5, Appl1	C 381	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 309	16	1.0	6008	1	US-08-005-002C-5	Sequence 5, Appl1	C 382	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 310	16	1.0	6008	1	US-08-487-203A-5	Sequence 5, Appl1	C 383	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 311	16	1.0	6619	3	US-08-683-007A-3	Sequence 3, Appl1	C 384	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 312	16	1.0	8133	2	US-08-480-604A-5	Sequence 5, Appl1	C 385	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 313	16	1.0	8133	2	US-08-480-604A-5	Sequence 5, Appl1	C 386	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 314	16	1.0	10898	4	US-08-481-658B-5	Sequence 5, Appl1	C 387	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 315	16	1.0	10898	4	US-08-477-504A-5	Sequence 5, Appl1	C 388	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 316	16	1.0	10898	4	US-08-467-56A-5	Sequence 5, Appl1	C 389	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 317	16	1.0	10898	4	US-08-485-762B-5	Sequence 5, Appl1	C 390	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 318	16	1.0	10898	5	US-08-787-739-5	Sequence 5, Appl1	C 391	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl
C 319	16	1.0	10898	5	US-08-487-077A-5	Sequence 5, Appl1	C 392	15	1.0	1000	6	PCT-US96-06385-96	Sequence 96, Appl

392	15	1.0	1924	3	US-08-756-317-1	Sequence 1, Appl1	C 465	15	1.0	3372	3	US-08-539-005-48	Sequence 48, Appl1
C 393	15	1.0	2030	2	US-08-530-950-1	Sequence 1, Appl1	466	15	1.0	3383	5	US-08-884-324-12	Sequence 12, Appl1
394	15	1.0	2060	1	US-07-721-761A-31	Sequence 31, Appl1	467	15	1.0	3440	5	US-08-471-791-27	Sequence 27, Appl1
395	15	1.0	2060	1	US-07-978-687-31	Sequence 31, Appl1	468	15	1.0	3440	6	PCT-US91-01746-27	Sequence 27, Appl1
396	15	1.0	2060	6	PCT-US91-01750-2	Sequence 2, Appl1	469	15	1.0	3465	2	US-08-578-649-3	Sequence 3, Appl1
397	15	1.0	2060	6	PCT-US91-05801-31	Sequence 31, Appl1	C 470	15	1.0	3625	2	US-08-356-354-3	Sequence 3, Appl1
398	15	1.0	2081	6	PCT-US91-01750-3	Sequence 3, Appl1	C 471	15	1.0	3625	2	US-08-778-656-3	Sequence 3, Appl1
C 399	15	1.0	2158	4	US-07-688-352C-3	Sequence 3, Appl1	C 472	15	1.0	3629	5	US-08-837-653-6	Sequence 6, Appl1
C 400	15	1.0	2158	4	US-08-474-379C-3	Sequence 3, Appl1	C 473	15	1.0	3704	4	US-08-702-153-1	Sequence 1, Appl1
C 401	15	1.0	2158	4	US-09-146-249A-3	Sequence 3, Appl1	C 474	15	1.0	3740	2	US-08-356-354-1	Sequence 1, Appl1
C 402	15	1.0	2158	6	PCT-US91-02714-3	Sequence 3, Appl1	C 475	15	1.0	3740	4	US-08-778-656-1	Sequence 1, Appl1
C 403	15	1.0	2173	4	US-08-036-210-6	Sequence 6, Appl1	C 476	15	1.0	3870	1	US-08-138-641-1	Sequence 1, Appl1
C 404	15	1.0	2173	4	US-08-449-609-6	Sequence 6, Appl1	C 477	15	1.0	3870	1	US-08-138-153-1	Sequence 1, Appl1
C 405	15	1.0	2245	3	US-08-617-801A-3	Sequence 3, Appl1	C 478	15	1.0	3871	4	US-08-599-455B-3	Sequence 3, Appl1
C 406	15	1.0	2248	1	US-08-369-780-1	Sequence 1, Appl1	C 479	15	1.0	3893	1	US-08-138-641-3	Sequence 3, Appl1
C 407	15	1.0	2248	1	US-08-475-682-1	Sequence 1, Appl1	C 480	15	1.0	3893	1	US-08-138-133-3	Sequence 3, Appl1
C 408	15	1.0	2248	2	US-08-780-833-1	Sequence 1, Appl1	C 481	15	1.0	3921	4	US-08-567-375-3	Sequence 3, Appl1
C 409	15	1.0	2248	2	US-08-636-036-1	Sequence 1, Appl1	C 482	15	1.0	3921	4	US-08-587-680A-3	Sequence 3, Appl1
C 410	15	1.0	2248	2	US-08-918-509-1	Sequence 1, Appl1	C 483	15	1.0	3973	4	US-08-036-210-21	Sequence 3, Appl1
C 411	15	1.0	2248	5	US-09-108-262-1	Sequence 1, Appl1	C 484	15	1.0	3973	4	US-08-449-609-21	Sequence 21, Appl1
C 412	15	1.0	2309	1	US-08-036-210-10	Sequence 10, Appl1	C 485	15	1.0	4002	1	US-08-331-488A-1	Sequence 1, Appl1
C 413	15	1.0	2309	4	US-08-449-609-10	Sequence 10, Appl1	C 486	15	1.0	4118	1	US-08-119-125A-3	Sequence 1, Appl1
C 414	15	1.0	2343	4	US-09-018-760-3	Sequence 3, Appl1	C 487	15	1.0	4291	3	US-08-417-210A-81	Sequence 3, Appl1
C 415	15	1.0	2344	5	US-08-893-852A-2	Sequence 2, Appl1	C 488	15	1.0	4406	1	US-08-233-788A-39	Sequence 39, Appl1
C 416	15	1.0	2369	3	US-08-523-742-1	Sequence 1, Appl1	C 489	15	1.0	4437	2	US-08-559-303B-72	Sequence 72, Appl1
C 417	15	1.0	2372	1	US-07-903-103-1	Sequence 1, Appl1	C 490	15	1.0	4454	4	US-08-738-172-2	Sequence 2, Appl1
C 418	15	1.0	2372	1	US-08-044-619A-1	Sequence 1, Appl1	C 491	15	1.0	4471	5	US-08-615-942A-1	Sequence 1, Appl1
C 419	15	1.0	2372	1	US-08-283-911-1	Sequence 1, Appl1	C 492	15	1.0	4818	5	US-08-817-926-27	Sequence 27, Appl1
C 420	15	1.0	2372	1	US-08-245-500A-2	Sequence 2, Appl1	C 493	15	1.0	4940	1	US-08-484-106-1	Sequence 1, Appl1
C 421	15	1.0	2372	1	US-08-390-546-2	Sequence 2, Appl1	C 494	15	1.0	4940	1	US-08-484-106-1	Sequence 1, Appl1
C 422	15	1.0	2372	1	US-08-390-479A-2	Sequence 2, Appl1	C 495	15	1.0	5057	3	US-08-365-486A-12	Sequence 12, Appl1
C 423	15	1.0	2372	2	US-08-557-393-2	Sequence 2, Appl1	C 496	15	1.0	5108	1	US-07-642-002-1	Sequence 1, Appl1
C 424	15	1.0	2372	2	US-08-390-516C-2	Sequence 2, Appl1	C 497	15	1.0	5285	4	US-08-609-049A-29	Sequence 29, Appl1
C 425	15	1.0	2372	2	US-08-390-517A-2	Sequence 2, Appl1	C 498	15	1.0	5893	2	US-08-592-126-54	Sequence 54, Appl1
C 426	15	1.0	2372	2	US-08-390-515A-2	Sequence 2, Appl1	C 499	15	1.0	5893	4	US-08-687-080-44	Sequence 44, Appl1
C 427	15	1.0	2372	5	US-08-801-718-2	Sequence 2, Appl1	C 500	15	1.0	5992	3	US-08-475-891A-3	Sequence 3, Appl1
C 428	15	1.0	2372	5	US-09-073-567-1	Sequence 1, Appl1	C 501	15	1.0	6063	1	US-08-195-744-4	Sequence 4, Appl1
C 429	15	1.0	2382	3	US-08-468-556-4	Sequence 4, Appl1	C 502	15	1.0	6063	3	US-08-788-279-4	Sequence 4, Appl1
C 430	15	1.0	2433	1	US-08-248-021A-1	Sequence 1, Appl1	C 503	15	1.0	6226	2	US-08-542-363-1	Sequence 1, Appl1
C 431	15	1.0	2440	1	US-08-160-861-2	Sequence 2, Appl1	C 504	15	1.0	6585	5	US-08-746-111-4	Sequence 4, Appl1
C 432	15	1.0	2442	2	US-08-542-363-3	Sequence 3, Appl1	C 505	15	1.0	6605	2	US-08-769-309A-4	Sequence 4, Appl1
C 433	15	1.0	2508	4	US-08-850-993-1	Sequence 1, Appl1	C 506	15	1.0	6914	2	US-08-920-812-22	Sequence 22, Appl1
C 434	15	1.0	2508	4	US-08-160-861-1	Sequence 1, Appl1	C 507	15	1.0	6914	2	US-08-920-827-22	Sequence 22, Appl1
C 435	15	1.0	2592	4	US-08-036-210-14	Sequence 14, Appl1	C 508	15	1.0	6914	2	US-08-921-177-22	Sequence 22, Appl1
C 436	15	1.0	2692	4	US-08-449-609-14	Sequence 14, Appl1	C 509	15	1.0	6914	2	US-08-362-577C-22	Sequence 22, Appl1
C 437	15	1.0	2791	4	US-09-016-000-8	Sequence 8, Appl1	C 510	15	1.0	6914	2	US-08-920-828-22	Sequence 22, Appl1
C 438	15	1.0	2877	3	US-08-693-697-35	Sequence 35, Appl1	C 511	15	1.0	7168	5	US-08-840-316-4	Sequence 5, Appl1
C 439	15	1.0	2880	3	US-08-693-697-32	Sequence 32, Appl1	C 512	15	1.0	7168	6	PCT-US93-0889A-4	Sequence 4, Appl1
C 440	15	1.0	2880	3	US-08-588-526-1	Sequence 1, Appl1	C 513	15	1.0	7168	6	PCT-US93-0889A-4	Sequence 4, Appl1
C 441	15	1.0	2936	4	US-08-338-172-1	Sequence 1, Appl1	C 514	15	1.0	7451	2	US-08-684-672-23	Sequence 23, Appl1
C 442	15	1.0	2974	1	US-08-408-486-2	Sequence 2, Appl1	C 515	15	1.0	7970	2	US-08-135-511-31	Sequence 31, Appl1
C 443	15	1.0	2980	3	US-08-463-081B-13	Sequence 13, Appl1	C 516	15	1.0	7977	2	US-08-187-453-31	Sequence 31, Appl1
C 444	15	1.0	2980	3	US-08-461-379A-13	Sequence 13, Appl1	C 517	15	1.0	7997	2	US-08-562-985A-7	Sequence 7, Appl1
C 445	15	1.0	2980	5	US-08-462-390B-13	Sequence 13, Appl1	C 518	15	1.0	10216	3	US-08-875-154-1	Sequence 1, Appl1
C 446	15	1.0	2980	5	US-08-463-074B-13	Sequence 13, Appl1	C 519	15	1.0	11464	5	US-08-884-324-13	Sequence 13, Appl1
C 447	15	1.0	2980	5	US-08-465-583C-13	Sequence 13, Appl1	C 520	15	1.0	11517	2	US-07-920-281C-1	Sequence 1, Appl1
C 448	15	1.0	2980	5	US-08-652-446-13	Sequence 13, Appl1	C 521	15	1.0	11601	3	US-08-222-617A-3	Sequence 3, Appl1
C 449	15	1.0	2991	2	US-08-355-888A-6	Sequence 6, Appl1	C 522	15	1.0	11601	3	US-08-222-617A-24	Sequence 24, Appl1
C 450	15	1.0	2991	3	US-08-588-190-1	Sequence 1, Appl1	C 523	15	1.0	13011	4	US-08-791-849A-14	Sequence 14, Appl1
C 451	15	1.0	2991	3	US-08-693-697-6	Sequence 6, Appl1	C 524	15	1.0	13999	4	US-08-444-644-24	Sequence 24, Appl1
C 452	15	1.0	2991	4	US-08-640-389A-1	Sequence 1, Appl1	C 525	15	1.0	12987	5	US-08-836-022A-10	Sequence 10, Appl1
C 453	15	1.0	3000	1	US-08-184-252A-1	Sequence 1, Appl1	C 526	15	1.0	22846	5	US-08-469-461-3	Sequence 3, Appl1
C 454	15	1.0	3000	6	PCT-US95-00601-1	Sequence 1, Appl1	C 527	15	1.0	22846	5	US-07-890-609-3	Sequence 3, Appl1
C 455	15	1.0	3000	6	PCT-US95-00601-1	Sequence 1, Appl1	C 528	15	1.0	26994	5	US-08-884-324-14	Sequence 14, Appl1
C 456	15	1.0	3060	5	US-08-098-141-1	Sequence 1, Appl1	C 529	15	1.0	246240	3	US-08-724-394A-20	Sequence 20, Appl1
C 457	15	1.0	3108	5	US-08-968-752B-3	Sequence 3, Appl1	C 530	15	1.0	246240	3	US-08-724-394A-21	Sequence 21, Appl1
C 458	15	1.0	3167	4	US-08-472-534-1	Sequence 1, Appl1	C 531	15	1.0	246240	3	US-08-724-394A-22	Sequence 22, Appl1
C 459	15	1.0	3222	5	US-08-968-752B-1	Sequence 1, Appl1	C 532	15	1.0	18	5	US-08-334-847-621	Sequence 621, App
C 460	15	1.0	3222	5	US-07-906-349A-1	Sequence 1, Appl1	C 533	15	1.0	18	5	US-08-411-796-403	Sequence 403, App
C 461	15	1.0	3372	1	US-08-167-035-1	Sequence 1, Appl1	C 534	15	1.0	18	5	US-08-471-039-403	Sequence 403, App
C 462	15	1.0	3372	1	US-08-167-035-48	Sequence 48, Appl1	C 535	15	1.0	21	6	PCT-US93-11198-403	Sequence 403, App
C 463	15	1.0	3372	1	US-08-208-887A-1	Sequence 1, Appl1	C 536	15	1.0	21	6	US-08-457-273B-40	Sequence 40, Appl1
C 464	15	1.0	3372	3	US-08-539-005-1	Sequence 1, Appl1	C 537	15	1.0	21	3	US-08-450-905B-146	Sequence 146, App

C 538	14	0.9	21	5	US-07-982-759F-146	Sequence 146, App	C 611	14	0.9	480	2	US-07-847-010-10	Sequence 10, App
539	14	0.9	24	4	US-08-913-477-5	Sequence 5, Appl	C 612	14	0.9	480	2	US-07-847-010-13	Sequence 13, Appl
540	14	0.9	24	4	US-08-913-477-5	Sequence 24, Appl	C 613	14	0.9	480	2	US-07-847-010-16	Sequence 16, Appl
541	14	0.9	27	1	US-08-220-606B-38	Sequence 38, Appl	C 614	14	0.9	480	2	US-07-847-010-19	Sequence 19, Appl
542	14	0.9	27	3	US-08-599-602-10	Sequence 10, Appl	C 615	14	0.9	483	2	US-08-592-126-118	Sequence 118, App
C 543	14	0.9	29	1	US-08-246-982A-14	Sequence 14, Appl	C 616	14	0.9	516	5	US-08-784-551C-1	Sequence 1, Appl
C 544	14	0.9	29	1	US-08-453-265-14	Sequence 14, Appl	617	14	0.9	525	1	US-08-434-411-1	Sequence 1, Appl
545	14	0.9	32	3	US-08-896-410-33	Sequence 33, Appl	618	14	0.9	525	2	US-08-434-402-1	Sequence 1, Appl
546	14	0.9	33	3	US-08-468-558-18	Sequence 18, Appl	619	14	0.9	525	2	US-08-783-288-1	Sequence 1, Appl
C 547	14	0.9	35	3	US-08-468-558-17	Sequence 17, Appl	620	14	0.9	525	4	US-08-890-640-1	Sequence 1, Appl
548	14	0.9	38	3	US-08-292-620A-2282	Sequence 2282, Ap	621	14	0.9	525	6	PCT-US94-12873-1	Sequence 1, Appl
C 549	14	0.9	42	2	US-08-261-660A-28	Sequence 28, Appl	622	14	0.9	525	7	5194592-25	Sequence 1, Appl
C 550	14	0.9	42	6	PCT-US94-06931-28	Sequence 28, Appl	623	14	0.9	537	6	PCT-US94-18873-16	Sequence 16, Appl
C 551	14	0.9	45	2	US-08-661-507-1	Sequence 1, Appl	624	14	0.9	546	3	US-08-797-689-13	Sequence 13, Appl
C 552	14	0.9	45	3	US-08-855-085-2	Sequence 2, Appl	625	14	0.9	546	5	US-08-469-118-177	Sequence 17, App
C 553	14	0.9	45	4	US-09-186-030-2	Sequence 2, Appl	626	14	0.9	546	5	US-08-469-118-178	Sequence 17, App
C 554	14	0.9	45	4	US-08-865-675-2	Sequence 2, Appl	627	14	0.9	546	5	US-08-469-118-179	Sequence 17, App
C 555	14	0.9	45	4	US-08-933-749-3	Sequence 3, Appl	628	14	0.9	546	6	PCT-US95-01185-177	Sequence 178, App
C 556	14	0.9	45	4	US-09-237-510-2	Sequence 2, Appl	629	14	0.9	546	6	PCT-US95-01185-178	Sequence 179, App
C 557	14	0.9	45	5	US-09-120-916-2	Sequence 2, Appl	630	14	0.9	555	6	PCT-US95-01185-179	Sequence 179, App
C 558	14	0.9	45	5	US-08-964-020-6	Sequence 6, Appl	631	14	0.9	555	6	PCT-US94-12873-13	Sequence 13, App
C 559	14	0.9	46	2	US-08-358-344B-8	Sequence 8, Appl	632	14	0.9	555	6	US-08-714-918-13	Sequence 13, Appl
C 560	14	0.9	46	2	US-08-358-344B-10	Sequence 10, Appl	633	14	0.9	594	5	US-08-742-023-3	Sequence 3, Appl
561	14	0.9	49	2	US-08-261-660A-27	Sequence 27, Appl	634	14	0.9	597	5	US-08-968-505-3	Sequence 3, Appl
C 562	14	0.9	49	3	US-08-652-558-20	Sequence 20, Appl	635	14	0.9	619	2	US-07-847-010-9	Sequence 9, Appl
C 563	14	0.9	49	3	US-08-652-558-21	Sequence 21, Appl	636	14	0.9	629	2	US-08-592-126-135	Sequence 135, App
C 564	14	0.9	49	3	US-08-417-210A-92	Sequence 92, Appl	637	14	0.9	655	2	US-07-847-010-15	Sequence 15, App
565	14	0.9	49	6	PCT-US94-06931-27	Sequence 27, Appl	638	14	0.9	677	5	US-08-718-804-3	Sequence 3, Appl
566	14	0.9	56	1	US-08-197-770-19	Sequence 19, Appl	639	14	0.9	677	6	PCT-US95-10973A-27	Sequence 27, Appl
C 567	14	0.9	56	1	US-08-049-264C-15	Sequence 15, Appl	640	14	0.9	692	2	US-08-465-293A-1	Sequence 1, Appl
C 568	14	0.9	56	1	US-08-476-562-15	Sequence 15, Appl	641	14	0.9	692	2	US-08-463-887A-1	Sequence 1, Appl
C 569	14	0.9	56	2	US-08-479-723A-15	Sequence 15, Appl	642	14	0.9	728	5	US-09-102-377-1	Sequence 1, Appl
C 570	14	0.9	56	6	PCT-US94-04310-15	Sequence 15, Appl	643	14	0.9	728	5	US-08-718-804-4	Sequence 4, Appl
C 571	14	0.9	58	7	5252474-1	Patent No. 5252474	644	14	0.9	728	6	PCT-US95-10973A-28	Sequence 28, Appl
C 572	14	0.9	72	3	US-08-599-602-5	Sequence 5, Appl	645	14	0.9	742	2	US-07-847-010-12	Sequence 12, Appl
C 573	14	0.9	120	4	US-08-484-575A-23	Sequence 23, Appl	646	14	0.9	750	4	US-08-687-080-74	Sequence 74, Appl
574	14	0.9	129	3	US-08-599-602-3	Sequence 3, Appl	647	14	0.9	777	3	US-08-618-011-3	Sequence 3, Appl
C 575	14	0.9	160	6	PCT-US93-06251-32	Sequence 32, Appl	648	14	0.9	777	3	US-08-618-011-5	Sequence 5, Appl
C 576	14	0.9	188	4	US-08-928-692-39	Sequence 39, Appl	649	14	0.9	795	3	US-08-898-779-1	Sequence 1, Appl
C 577	14	0.9	200	6	PCT-US94-09450-10	Sequence 10, Appl	650	14	0.9	795	3	US-08-898-779-3	Sequence 3, Appl
C 578	14	0.9	225	1	US-08-264-534-14	Sequence 24, Appl	651	14	0.9	816	4	US-08-822-028-25	Sequence 25, Appl
C 579	14	0.9	225	2	US-08-465-500-24	Sequence 24, Appl	652	14	0.9	816	4	US-08-822-028-26	Sequence 26, Appl
C 580	14	0.9	225	3	US-08-346-126-24	Sequence 24, Appl	653	14	0.9	830	2	US-08-822-028-25	Sequence 25, Appl
C 581	14	0.9	225	3	US-08-346-128-24	Sequence 24, Appl	654	14	0.9	853	1	US-07-876-284-1	Sequence 1, Appl
C 582	14	0.9	255	2	US-08-680-726A-83	Sequence 83, Appl	655	14	0.9	860	2	US-07-847-010-18	Sequence 18, Appl
C 583	14	0.9	295	1	US-08-264-534-15	Sequence 15, Appl	656	14	0.9	864	4	US-08-924-759-15	Sequence 15, Appl
C 584	14	0.9	295	2	US-08-083-590A-7	Sequence 7, Appl	657	14	0.9	864	4	US-08-822-028-5	Sequence 5, Appl
C 585	14	0.9	295	2	US-08-465-500-15	Sequence 15, Appl	658	14	0.9	864	4	US-08-595-668C-11	Sequence 11, Appl
C 586	14	0.9	295	3	US-08-346-126-15	Sequence 15, Appl	659	14	0.9	865	4	US-08-822-028-3	Sequence 3, Appl
C 587	14	0.9	318	1	US-08-171-385-3	Sequence 3, Appl	660	14	0.9	871	5	PCT-US91-06234-7	Sequence 7, Appl
588	14	0.9	318	1	US-08-361-441B-3	Sequence 3, Appl	661	14	0.9	906	6	US-08-469-118-53	Sequence 53, Appl
589	14	0.9	400	5	US-08-714-918-23	Sequence 23, Appl	662	14	0.9	906	6	PCT-US95-01185-53	Sequence 53, Appl
590	14	0.9	423	1	US-08-470-179-77	Sequence 77, Appl	663	14	0.9	921	5	US-08-469-118-56	Sequence 56, Appl
C 591	14	0.9	423	1	US-08-470-179-134	Sequence 134, App	664	14	0.9	921	5	US-08-469-118-59	Sequence 59, Appl
C 592	14	0.9	444	2	US-08-688-609-2	Sequence 2, Appl	665	14	0.9	921	5	US-08-469-118-60	Sequence 60, Appl
C 593	14	0.9	461	3	US-08-977-554-3	Sequence 3, Appl	666	14	0.9	921	5	US-08-469-118-72	Sequence 72, Appl
594	14	0.9	470	1	US-07-998-003A-66	Sequence 66, Appl	667	14	0.9	921	5	US-08-469-118-75	Sequence 75, Appl
595	14	0.9	470	1	US-08-453-274B-66	Sequence 66, Appl	668	14	0.9	921	5	US-08-469-118-80	Sequence 80, Appl
596	14	0.9	470	1	US-08-453-274B-66	Sequence 66, Appl	669	14	0.9	921	5	US-08-469-118-84	Sequence 84, Appl
597	14	0.9	470	2	US-08-268-161A-66	Sequence 66, Appl	670	14	0.9	921	5	US-08-469-118-85	Sequence 85, Appl
598	14	0.9	470	2	US-08-453-702A-66	Sequence 66, Appl	671	14	0.9	921	5	US-08-469-118-87	Sequence 87, Appl
599	14	0.9	470	6	PCT-US93-12588-66	Sequence 66, Appl	672	14	0.9	921	5	US-08-469-118-85	Sequence 85, Appl
600	14	0.9	470	6	PCT-US93-12588-66	Sequence 66, Appl	673	14	0.9	921	5	US-08-469-118-87	Sequence 87, Appl
601	14	0.9	470	6	PCT-US93-08071-66	Sequence 66, Appl	674	14	0.9	921	6	PCT-US93-01185-56	Sequence 56, Appl
602	14	0.9	471	3	US-08-599-602-1	Sequence 1, Appl	675	14	0.9	921	6	PCT-US95-01185-59	Sequence 59, Appl
603	14	0.9	473	2	US-08-480-784-23	Sequence 23, Appl	676	14	0.9	921	6	PCT-US95-01185-60	Sequence 60, Appl
604	14	0.9	473	2	US-08-483-553-23	Sequence 23, Appl	677	14	0.9	921	6	PCT-US95-01185-72	Sequence 72, Appl
605	14	0.9	473	2	US-08-487-002-23	Sequence 23, Appl	678	14	0.9	921	6	PCT-US93-01185-75	Sequence 75, Appl
606	14	0.9	473	2	US-08-483-554B-23	Sequence 23, Appl	679	14	0.9	921	6	PCT-US93-01185-80	Sequence 80, Appl
607	14	0.9	473	2	US-08-488-011B-23	Sequence 23, Appl	680	14	0.9	921	6	PCT-US95-01185-84	Sequence 84, Appl
608	14	0.9	473	6	PCT-US95-10202-23	Sequence 23, Appl	681	14	0.9	921	6	PCT-US95-01185-85	Sequence 85, Appl
609	14	0.9	473	6	PCT-US95-10203-23	Sequence 23, Appl	682	14	0.9	921	6	PCT-US95-01185-87	Sequence 87, Appl
610	14	0.9	473	6	PCT-US95-10220-23	Sequence 23, Appl	683	14	0.9	924	1	US-07-593-657-4	Sequence 4, Appl

C 684	14	0.9	950	5	US-08-581-148C-10	Sequence 10, Appl	757	14	0.9	1299	2	US-08-463-262A-1	Sequence 1, Appl1
C 685	14	0.9	966	5	US-08-469-318-68	Sequence 68, Appl	758	14	0.9	1299	2	US-08-463-989-1	Sequence 1, Appl1
C 686	14	0.9	966	5	US-08-469-318-70	Sequence 70, Appl	C 759	14	0.9	1391	6	PCT-US96-1051-19	Sequence 19, Appl
C 687	14	0.9	966	5	US-08-469-318-71	Sequence 71, Appl	C 760	14	0.9	1309	4	US-08-933-750C-63	Sequence 63, Appl
C 688	14	0.9	966	5	US-08-469-318-73	Sequence 73, Appl	C 761	14	0.9	1329	1	US-08-278-630A-8	Sequence 8, Appl
C 689	14	0.9	966	5	US-08-469-318-77	Sequence 77, Appl	C 762	14	0.9	1330	3	US-08-807-200-1	Sequence 1, Appl1
C 690	14	0.9	966	5	US-08-469-318-79	Sequence 79, Appl	C 763	14	0.9	1345	4	US-08-702-153-3	Sequence 3, Appl1
C 691	14	0.9	966	5	US-08-469-318-81	Sequence 81, Appl	C 764	14	0.9	1350	1	US-07-792-466-7	Sequence 7, Appl1
C 692	14	0.9	966	5	US-08-469-318-89	Sequence 89, Appl	C 765	14	0.9	1350	2	US-08-474-633A-1	Sequence 1, Appl1
C 693	14	0.9	966	6	PCT-US95-01185-68	Sequence 68, Appl	C 766	14	0.9	1350	5	US-08-737-552B-4	Sequence 4, Appl1
C 694	14	0.9	966	6	PCT-US95-01185-70	Sequence 70, Appl	C 767	14	0.9	1350	6	PCT-US93-02480-1	Sequence 1, Appl1
C 695	14	0.9	966	6	PCT-US95-01185-71	Sequence 71, Appl	C 768	14	0.9	1355	3	US-08-618-408B-3	Sequence 3, Appl1
C 696	14	0.9	966	6	PCT-US95-01185-73	Sequence 73, Appl	C 769	14	0.9	1356	3	US-08-484-126-4	Sequence 4, Appl1
C 697	14	0.9	966	6	PCT-US95-01185-77	Sequence 77, Appl	C 770	14	0.9	1373	6	PCT-US92-02977-6	Sequence 6, Appl1
C 698	14	0.9	966	6	PCT-US95-01185-79	Sequence 79, Appl	C 771	14	0.9	1373	6	PCT-US95-03032-2	Sequence 3, Appl1
C 699	14	0.9	966	6	PCT-US95-01185-81	Sequence 81, Appl	C 772	14	0.9	1383	1	US-08-289-709-2	Sequence 2, Appl1
C 700	14	0.9	966	6	PCT-US95-01185-89	Sequence 89, Appl	C 773	14	0.9	1383	1	US-08-602-656-2	Sequence 2, Appl1
C 701	14	0.9	982	4	US-08-822-028-1	Sequence 1, Appl1	C 774	14	0.9	1389	2	US-08-702-344-20	Sequence 20, Appl
C 702	14	0.9	984	5	US-08-469-318-83	Sequence 83, Appl	C 775	14	0.9	1395	1	US-07-991-867B-25	Sequence 25, Appl
C 703	14	0.9	984	6	PCT-US95-01185-83	Sequence 83, Appl	C 776	14	0.9	1395	1	US-07-991-867B-25	Sequence 25, Appl
C 704	14	0.9	1005	5	US-09-103-359-4	Sequence 4, Appl1	C 777	14	0.9	1395	2	US-08-107-755A-25	Sequence 25, Appl
C 705	14	0.9	1017	5	US-08-469-318-67	Sequence 67, Appl	C 778	14	0.9	1395	2	US-08-107-755A-25	Sequence 25, Appl
C 706	14	0.9	1017	6	PCT-US95-01185-67	Sequence 67, Appl	C 779	14	0.9	1395	4	US-08-544-332-25	Sequence 25, Appl
C 707	14	0.9	1018	7	5187262-7	Sequence 67, Appl	C 780	14	0.9	1395	4	US-08-544-332-25	Sequence 25, Appl
C 708	14	0.9	1047	5	US-08-469-318-65	Sequence 65, Appl	C 781	14	0.9	1395	4	US-08-544-332-25	Sequence 25, Appl
C 709	14	0.9	1047	5	US-08-469-318-74	Sequence 74, Appl	C 782	14	0.9	1400	3	US-08-305-764C-59	Sequence 59, Appl
C 710	14	0.9	1047	5	US-08-469-318-76	Sequence 76, Appl	C 783	14	0.9	1425	1	US-07-715-184-3	Sequence 3, Appl1
C 711	14	0.9	1047	6	PCT-US95-01185-65	Sequence 65, Appl	C 784	14	0.9	1425	1	US-07-876-280-6	Sequence 6, Appl1
C 712	14	0.9	1047	6	PCT-US95-01185-74	Sequence 74, Appl	C 785	14	0.9	1425	1	US-07-876-280-6	Sequence 6, Appl1
C 713	14	0.9	1047	6	PCT-US95-01185-76	Sequence 76, Appl	C 786	14	0.9	1425	1	US-07-935-310A-1	Sequence 27, Appl
C 714	14	0.9	1057	2	US-08-470-202-45	Sequence 45, Appl	C 787	14	0.9	1425	1	US-07-935-310A-1	Sequence 27, Appl
C 715	14	0.9	1057	2	US-08-471-770-45	Sequence 45, Appl	C 788	14	0.9	1425	1	US-07-828-788A-1	Sequence 1, Appl1
C 716	14	0.9	1057	3	US-08-468-059-45	Sequence 45, Appl	C 789	14	0.9	1425	1	US-08-049-783-3	Sequence 3, Appl1
C 717	14	0.9	1062	1	US-08-426-169-6	Sequence 6, Appl1	C 790	14	0.9	1425	2	US-08-147-189-1	Sequence 1, Appl1
C 718	14	0.9	1062	6	PCT-US95-09470-6	Sequence 6, Appl1	C 791	14	0.9	1425	2	US-08-316-301A-7	Sequence 7, Appl1
C 719	14	0.9	1080	7	5474933-1	Sequence 6, Appl1	C 792	14	0.9	1425	3	US-08-904-278-1	Sequence 7, Appl1
C 720	14	0.9	1083	6	PCT-US94-09450-7	Sequence 7, Appl1	C 793	14	0.9	1425	6	PCT-US92-03624-7	Sequence 7, Appl1
C 721	14	0.9	1121	2	US-08-523-855A-2	Sequence 2, Appl1	C 794	14	0.9	1425	6	PCT-US92-11337-1	Sequence 1, Appl1
C 722	14	0.9	1121	7	5256558-7	Sequence 7, Appl1	C 795	14	0.9	1434	3	PCT-US93-07409-1	Sequence 1, Appl1
C 723	14	0.9	1124	4	US-08-578-551-1	Sequence 1, Appl1	C 796	14	0.9	1440	3	US-08-476-254-1	Sequence 1, Appl1
C 724	14	0.9	1124	4	US-09-190-982-1	Sequence 1, Appl1	C 797	14	0.9	1440	5	US-08-807-200-11	Sequence 11, Appl
C 725	14	0.9	1126	6	PCT-US95-17083-16	Sequence 16, Appl	C 798	14	0.9	1443	6	US-08-852-782A-2	Sequence 2, Appl1
C 726	14	0.9	1182	3	US-08-977-554-1	Sequence 1, Appl1	C 799	14	0.9	1463	6	PCT-US96-10521-33	Sequence 33, Appl
C 727	14	0.9	1188	3	US-08-921-426-6	Sequence 6, Appl1	C 800	14	0.9	1499	1	PCT-US93-05640-10	Sequence 10, Appl
C 728	14	0.9	1188	5	US-08-816-915-6	Sequence 6, Appl1	C 801	14	0.9	1499	4	US-08-415-818-11	Sequence 11, Appl
C 729	14	0.9	1188	6	PCT-US95-07743-6	Sequence 6, Appl1	C 802	14	0.9	1499	4	US-08-894-236-11	Sequence 11, Appl
C 730	14	0.9	1189	4	US-08-450-042A-2	Sequence 2, Appl1	C 803	14	0.9	1506	6	PCT-US96-01444-11	Sequence 11, Appl
C 731	14	0.9	1195	7	5240848-6	Sequence 2, Appl1	C 804	14	0.9	1506	1	US-08-149-105-2	Sequence 1, Appl
C 732	14	0.9	1200	3	US-08-642-045B-3	Sequence 3, Appl1	C 805	14	0.9	1512	2	US-08-317-847-2	Sequence 2, Appl1
C 733	14	0.9	1201	1	US-08-181-271A-26	Sequence 26, Appl	C 806	14	0.9	1512	2	US-08-594-808B-6	Sequence 6, Appl1
C 734	14	0.9	1201	1	US-08-449-315-26	Sequence 26, Appl	C 807	14	0.9	1524	2	US-08-409-122-1	Sequence 1, Appl1
C 735	14	0.9	1201	1	US-08-444-803-26	Sequence 26, Appl	C 808	14	0.9	1524	2	US-08-408-666-1	Sequence 1, Appl1
C 736	14	0.9	1201	1	US-08-449-043-26	Sequence 26, Appl	C 809	14	0.9	1541	3	US-08-668-289-9	Sequence 9, Appl1
C 737	14	0.9	1201	2	US-08-456-265A-26	Sequence 26, Appl	C 810	14	0.9	1541	3	US-09-167-112-9	Sequence 9, Appl1
C 738	14	0.9	1201	2	US-08-456-265A-26	Sequence 26, Appl	C 811	14	0.9	1550	5	US-08-912-560-1	Sequence 1, Appl1
C 739	14	0.9	1201	2	US-08-455-416-26	Sequence 26, Appl	C 812	14	0.9	1552	5	US-08-948-705-1	Sequence 1, Appl1
C 740	14	0.9	1201	2	US-08-455-244-26	Sequence 26, Appl	C 813	14	0.9	1581	4	US-08-853-652A-15	Sequence 15, Appl
C 741	14	0.9	1201	3	US-08-457-364-26	Sequence 26, Appl	C 814	14	0.9	1592	1	US-08-044-618-4	Sequence 4, Appl1
C 742	14	0.9	1201	3	US-08-456-262-26	Sequence 26, Appl	C 815	14	0.9	1593	1	US-08-307-499-50	Sequence 50, Appl
C 743	14	0.9	1201	3	US-08-456-240-26	Sequence 26, Appl	C 816	14	0.9	1603	1	US-07-885-970A-19	Sequence 19, Appl
C 744	14	0.9	1201	3	US-08-455-736-26	Sequence 26, Appl	C 817	14	0.9	1603	1	US-08-298-687A-19	Sequence 19, Appl
C 745	14	0.9	1201	4	US-08-971-217-26	Sequence 26, Appl	C 818	14	0.9	1644	5	US-08-852-782A-1	Sequence 1, Appl1
C 746	14	0.9	1216	4	US-08-078-311-20	Sequence 20, Appl	C 819	14	0.9	1674	2	US-08-852-782A-1	Sequence 1, Appl1
C 747	14	0.9	1216	4	US-08-460-402-20	Sequence 20, Appl	C 820	14	0.9	1686	2	US-08-322-493-1	Sequence 1, Appl1
C 748	14	0.9	1230	2	US-08-700-442A-8	Sequence 8, Appl1	C 821	14	0.9	1732	2	US-08-480-346-2	Sequence 2, Appl1
C 749	14	0.9	1242	2	US-07-877-516-1	Sequence 1, Appl1	C 822	14	0.9	1732	2	US-08-480-346-2	Sequence 2, Appl1
C 750	14	0.9	1242	6	PCT-US93-03077-4	Sequence 4, Appl1	C 823	14	0.9	1732	3	US-08-243-541-2	Sequence 2, Appl1
C 751	14	0.9	1248	2	US-08-602-010A-9	Sequence 9, Appl1	C 824	14	0.9	1743	1	US-08-171-382-3	Sequence 3, Appl1
C 752	14	0.9	1248	2	US-08-680-726A-9	Sequence 9, Appl1	C 825	14	0.9	1743	1	US-08-171-382-3	Sequence 3, Appl1
C 753	14	0.9	1263	6	PCT-US96-10602-9	Sequence 9, Appl1	C 826	14	0.9	1743	1	US-08-309-420-3	Sequence 3, Appl1
C 754	14	0.9	1278	4	US-08-765-268A-2	Sequence 2, Appl1	C 827	14	0.9	1743	1	US-08-309-420-3	Sequence 3, Appl1
C 755	14	0.9	1286	3	US-08-807-200-3	Sequence 3, Appl1	C 828	14	0.9	1743	6	US-08-841-178-20	Sequence 20, Appl
C 756	14	0.9	1294	2	US-08-464-523B-3	Sequence 3, Appl1	C 829	14	0.9	1743	6	PCT-US95-11878-3	Sequence 3, Appl1

C 830	14	0.9	1752	1	US-07-629-1041-1	Sequence 1, Appl1	C 903	14	0.9	2278	2	US-08-526-813-1	Sequence 1, Appl1
C 831	14	0.9	1767	1	US-08-399-646-1	Sequence 1, Appl1	C 904	14	0.9	2278	6	PCT-US95-08554-1	Sequence 1, Appl1
C 832	14	0.9	1767	2	US-08-607-321-1	Sequence 1, Appl1	C 905	14	0.9	2314	5	US-08-105-454-1	Sequence 1, Appl1
C 833	14	0.9	1767	3	US-08-961-240-1	Sequence 1, Appl1	C 906	14	0.9	2322	2	US-08-618-164-1	Sequence 1, Appl1
C 834	14	0.9	1767	3	US-08-605-501-1	Sequence 1, Appl1	C 907	14	0.9	2325	1	US-08-314-309A-20	Sequence 20, Appl1
C 835	14	0.9	1769	2	US-08-233-155A-19	Sequence 19, Appl1	C 908	14	0.9	2336	5	US-08-389-564B-2	Sequence 2, Appl1
C 836	14	0.9	1769	4	US-08-765-268A-1	Sequence 1, Appl1	C 909	14	0.9	2336	4	US-08-466-047B-2	Sequence 2, Appl1
C 837	14	0.9	1776	2	US-08-464-523B-4	Sequence 4, Appl1	C 910	14	0.9	2363	5	US-08-671-978A-11	Sequence 11, Appl1
C 838	14	0.9	1778	6	PCT-US91-02958-9	Sequence 9, Appl1	C 911	14	0.9	2382	1	US-08-256-938-1	Sequence 1, Appl1
C 839	14	0.9	1780	5	PCT-US94-12913A-17	Sequence 17, Appl1	C 912	14	0.9	2392	1	US-08-171-382-5	Sequence 5, Appl1
C 840	14	0.9	1789	5	US-08-948-705-6	Sequence 6, Appl1	C 913	14	0.9	2392	1	US-08-309-420-5	Sequence 5, Appl1
C 841	14	0.9	1823	2	US-08-680-726A-77	Sequence 77, Appl1	C 914	14	0.9	2392	1	US-08-309-419-5	Sequence 5, Appl1
C 842	14	0.9	1823	3	US-08-680-726A-78	Sequence 78, Appl1	C 915	14	0.9	2392	6	PCT-US95-11856-5	Sequence 5, Appl1
C 843	14	0.9	1840	3	US-08-484-993B-15	Sequence 15, Appl1	C 916	14	0.9	2392	6	PCT-US95-11878-5	Sequence 5, Appl1
C 844	14	0.9	1840	4	US-08-484-158B-15	Sequence 15, Appl1	C 917	14	0.9	2401	2	US-08-480-346-1	Sequence 1, Appl1
C 845	14	0.9	1840	4	US-08-484-596A-15	Sequence 15, Appl1	C 918	14	0.9	2401	3	US-08-243-541-1	Sequence 1, Appl1
C 846	14	0.9	1840	4	US-08-480-150A-15	Sequence 15, Appl1	C 919	14	0.9	2401	3	US-08-480-344-1	Sequence 1, Appl1
C 847	14	0.9	1840	5	US-08-458-731-15	Sequence 15, Appl1	C 920	14	0.9	2409	1	US-07-911-531-18	Sequence 18, Appl1
C 848	14	0.9	1840	5	US-08-149-223A-15	Sequence 15, Appl1	C 921	14	0.9	2409	1	US-07-693-636A-18	Sequence 18, Appl1
C 849	14	0.9	1867	3	US-08-516-801-1	Sequence 1, Appl1	C 922	14	0.9	2417	2	US-08-011-398B-1	Sequence 1, Appl1
C 850	14	0.9	1867	6	PCT-US95-06683-1	Sequence 49, Appl1	C 923	14	0.9	2417	2	US-08-464-051-1	Sequence 1, Appl1
C 851	14	0.9	1880	1	US-08-247-475-49	Sequence 49, Appl1	C 924	14	0.9	2417	3	US-08-462-488-1	Sequence 1, Appl1
C 852	14	0.9	1880	1	US-08-479-650-48	Sequence 49, Appl1	C 925	14	0.9	2417	5	US-08-554-385-2	Sequence 2, Appl1
C 853	14	0.9	1880	2	US-08-191-866D-80	Sequence 80, Appl1	C 926	14	0.9	2427	2	US-08-490-099-1	Sequence 1, Appl1
C 854	14	0.9	1880	2	US-08-674-169-49	Sequence 49, Appl1	C 927	14	0.9	2430	3	US-08-062-368-1	Sequence 1, Appl1
C 855	14	0.9	1880	3	US-08-185-949B-80	Sequence 80, Appl1	C 928	14	0.9	2430	3	US-08-820-170A-35	Sequence 35, Appl1
C 856	14	0.9	1883	2	US-08-665-220-3	Sequence 3, Appl1	C 929	14	0.9	2436	2	US-08-306-601B-16	Sequence 16, Appl1
C 857	14	0.9	1911	2	US-08-258-420-7	Sequence 8, Appl1	C 930	14	0.9	2448	4	US-08-687-080-111	Sequence 11, App
C 858	14	0.9	1914	2	US-08-258-420-7	Sequence 7, Appl1	C 931	14	0.9	2451	3	US-08-820-170A-29	Sequence 29, Appl1
C 859	14	0.9	1930	1	US-07-982-112-1	Sequence 1, Appl1	C 932	14	0.9	2455	1	US-08-256-938-3	Sequence 3, Appl1
C 860	14	0.9	1939	7	US-07-982-112-1	Patent No. 5198542-3	C 933	14	0.9	2455	1	US-08-797-689-15	Sequence 15, Appl1
C 861	14	0.9	1942	5	US-08-627-907A-3	Sequence 6, Appl1	C 934	14	0.9	2471	1	US-08-444-231-18	Sequence 18, Appl1
C 862	14	0.9	2009	2	US-08-201-118-6	Sequence 6, Appl1	C 935	14	0.9	2471	1	US-08-152-443A-18	Sequence 18, Appl1
C 863	14	0.9	2009	4	US-08-238-821B-6	Sequence 6, Appl1	C 936	14	0.9	2472	2	US-08-425-289A-2	Sequence 2, Appl1
C 864	14	0.9	2009	6	PCT-US93-05744-6	Sequence 6, Appl1	C 937	14	0.9	2474	4	US-08-666-352A-2	Sequence 2, Appl1
C 865	14	0.9	2014	3	US-08-525-742-7	Sequence 7, Appl1	C 938	14	0.9	2487	3	US-08-820-170A-32	Sequence 32, Appl1
C 866	14	0.9	2020	3	US-08-417-210A-136	Sequence 136, App	C 939	14	0.9	2505	3	US-08-890-094-1	Sequence 1, Appl1
C 867	14	0.9	2030	4	US-08-923-536A-1	Sequence 1, Appl1	C 940	14	0.9	2517	1	US-07-906-930E-1	Sequence 1, Appl1
C 868	14	0.9	2049	3	US-08-417-210A-117	Sequence 117, App	C 941	14	0.9	2529	3	US-08-867-129-1	Sequence 1, Appl1
C 869	14	0.9	2070	3	US-08-890-094-17	Sequence 17, Appl1	C 942	14	0.9	2534	3	US-08-219-237B-1	Sequence 1, Appl1
C 870	14	0.9	2071	2	US-08-393-985-15	Sequence 15, Appl1	C 943	14	0.9	2550	7	5258287-23	Patent No. 5258287
C 871	14	0.9	2074	1	US-08-414-685-1	Sequence 1, Appl1	C 944	14	0.9	2559	3	US-08-248-839C-10	Sequence 10, Appl1
C 872	14	0.9	2095	1	US-08-261-086-1	Sequence 1, Appl1	C 945	14	0.9	2587	3	US-08-326-286-3	Sequence 3, Appl1
C 873	14	0.9	2095	1	US-08-305-764C-55	Sequence 55, Appl1	C 946	14	0.9	2619	6	PCT-US96-10521-17	Sequence 17, Appl1
C 874	14	0.9	2115	1	US-08-329-681A-3	Sequence 3, Appl1	C 947	14	0.9	2629	4	US-09-012-030-1	Sequence 1, Appl1
C 875	14	0.9	2126	2	US-08-463-262A-2	Sequence 2, Appl1	C 948	14	0.9	2629	4	US-08-590-454-1	Sequence 1, Appl1
C 876	14	0.9	2126	2	US-08-463-985-2	Sequence 2, Appl1	C 949	14	0.9	2631	2	US-08-208-036-13	Sequence 13, Appl1
C 877	14	0.9	2147	1	US-08-256-135-1	Sequence 1, Appl1	C 950	14	0.9	2631	2	US-08-428-823-13	Sequence 13, Appl1
C 878	14	0.9	2147	4	US-08-950-737-1	Sequence 7, Appl1	C 951	14	0.9	2649	4	US-08-718-964A-1	Sequence 1, Appl1
C 879	14	0.9	2147	4	US-08-973-461A-7	Sequence 7, Appl1	C 952	14	0.9	2649	4	US-09-050-954A-1	Sequence 1, Appl1
C 880	14	0.9	2147	5	US-08-648-010-7	Sequence 7, Appl1	C 953	14	0.9	2649	4	US-08-842-341-1	Sequence 1, Appl1
C 881	14	0.9	2150	1	US-08-198-446B-10	Sequence 10, Appl1	C 954	14	0.9	2661	2	US-08-351-413-1	Sequence 1, Appl1
C 882	14	0.9	2150	3	US-08-870-693-10	Sequence 10, Appl1	C 955	14	0.9	2661	2	US-09-025-583-1	Sequence 1, Appl1
C 883	14	0.9	2152	1	US-07-779-890-1	Sequence 1, Appl1	C 956	14	0.9	2682	4	US-08-044-618-3	Sequence 3, Appl1
C 884	14	0.9	2152	2	US-07-779-890-1	Sequence 1, Appl1	C 957	14	0.9	2682	4	US-08-867-941-7	Sequence 7, Appl1
C 885	14	0.9	2152	6	PCT-US93-05640-1	Sequence 1, Appl1	C 958	14	0.9	2689	4	US-08-876-546A-15	Sequence 15, Appl1
C 886	14	0.9	2152	6	US-08-399-646-11	Sequence 11, Appl1	C 959	14	0.9	2692	1	US-07-932-454A-2	Sequence 2, Appl1
C 887	14	0.9	2161	2	US-08-607-321-11	Sequence 11, Appl1	C 960	14	0.9	2694	1	US-08-867-941-2	Sequence 2, Appl1
C 888	14	0.9	2161	3	US-08-961-240-11	Sequence 11, Appl1	C 961	14	0.9	2733	2	US-08-676-967-3	Sequence 3, Appl1
C 889	14	0.9	2161	3	US-08-605-501-11	Sequence 11, Appl1	C 962	14	0.9	2733	2	US-08-676-974-3	Sequence 3, Appl1
C 890	14	0.9	2180	5	US-09-003-217-1	Sequence 1, Appl1	C 963	14	0.9	2733	1	US-09-098-487-3	Sequence 3, Appl1
C 891	14	0.9	2185	4	US-08-714-918-83	Sequence 83, Appl1	C 964	14	0.9	2787	1	US-08-530-010-2	Sequence 2, Appl1
C 892	14	0.9	2188	1	US-07-882-925A-4	Sequence 4, Appl1	C 965	14	0.9	2787	1	US-08-530-010-4	Sequence 4, Appl1
C 893	14	0.9	2188	1	US-08-184-012C-4	Sequence 4, Appl1	C 966	14	0.9	2787	1	US-08-530-010-6	Sequence 6, Appl1
C 894	14	0.9	2200	2	US-08-272-235-21	Sequence 21, Appl1	C 967	14	0.9	2787	1	US-08-530-010-8	Sequence 8, Appl1
C 895	14	0.9	2200	6	PCT-US95-08565-21	Sequence 21, Appl1	C 968	14	0.9	2787	2	US-08-530-010-10	Sequence 10, Appl1
C 896	14	0.9	2219	5	US-08-510-646B-17	Sequence 17, Appl1	C 969	14	0.9	2787	2	US-08-484-101B-2	Sequence 2, Appl1
C 897	14	0.9	2234	5	US-08-714-918-75	Sequence 75, Appl1	C 970	14	0.9	2787	2	US-08-484-101B-6	Sequence 6, Appl1
C 898	14	0.9	2234	5	US-08-201-118-12	Sequence 12, Appl1	C 971	14	0.9	2787	2	US-08-484-101B-8	Sequence 8, Appl1
C 899	14	0.9	2258	4	US-08-238-821B-12	Sequence 12, Appl1	C 972	14	0.9	2787	2	US-08-484-101B-8	Sequence 8, Appl1
C 900	14	0.9	2258	6	PCT-US93-05744-12	Sequence 12, Appl1	C 973	14	0.9	2787	2	US-08-484-101B-10	Sequence 10, Appl1
C 901	14	0.9	2268	3	US-08-873-093-2	Sequence 2, Appl1	C 974	14	0.9	2815	1	US-07-671-817A-2	Sequence 2, Appl1
C 902	14	0.9	2278	1	US-08-258-188-1	Sequence 1, Appl1	C 975	14	0.9	2817	6	PCT-US93-05444-1	Sequence 1, Appl1

976 14 0.9 2834 1 US-08-276-151-8 Sequence 8, Appl1
 C 977 14 0.9 2859 3 US-08-637-7638-7 Sequence 7, Appl1
 C 978 14 0.9 2859 5 US-09-170-354-7 Sequence 7, Appl1
 979 14 0.9 2875 1 US-08-328-314-1 Sequence 1, Appl1
 980 14 0.9 2875 2 US-08-731-045-1 Sequence 1, Appl1
 981 14 0.9 2882 1 US-08-424-567-1 Sequence 1, Appl1
 982 14 0.9 2882 3 US-08-711-928-1 Sequence 1, Appl1
 C 983 14 0.9 2887 6 PCT-US96-10521-14 Sequence 14, Appl1
 984 14 0.9 2903 3 US-08-310-912A-1 Sequence 1, Appl1
 985 14 0.9 2903 6 PCT-US95-04570-1 Sequence 1, Appl1
 986 14 0.9 2903 6 PCT-US95-04589-1 Sequence 1, Appl1
 987 14 0.9 2939 1 US-07-806-930E-3 Sequence 3, Appl1
 C 988 14 0.9 2964 1 US-08-286-819A-18 Sequence 18, Appl1
 C 989 14 0.9 2964 5 US-08-980-357-18 Sequence 18, Appl1
 990 14 0.9 2977 3 US-08-820-170A-36 Sequence 36, Appl1
 991 14 0.9 2990 1 US-07-671-817A-1 Sequence 1, Appl1
 C 992 14 0.9 3003 7 5182260-18 Patent No. 5182260
 993 14 0.9 3012 1 US-07-991-867B-21 Sequence 21, Appl1
 994 14 0.9 3012 2 US-08-107-755A-21 Sequence 21, Appl1
 C 995 14 0.9 3012 4 US-08-544-332-21 Sequence 21, Appl1
 C 996 14 0.9 3035 6 PCT-US94-09235-1 Sequence 1, Appl1
 997 14 0.9 3063 3 US-08-184-009-169 Sequence 169, App
 998 14 0.9 3063 4 US-08-458-356-169 Sequence 3, Appl1
 999 14 0.9 3066 1 US-07-671-817A-3 Sequence 3, Appl1
 1000 14 0.9 3070 1 US-08-428-732-3 Sequence 3, Appl1

ALIGNMENTS

RESULT 1
 US-08-913-477-1
 : Sequence 1, Application US/08913477
 : Patent No. 5985285
 : GENERAL INFORMATION:
 : APPLICANT: Tibball, Richard W.
 : APPLICANT: Williamson, Ethel D.
 : APPLICANT: Leary, Sophie E.C.
 : APPLICANT: Oyston, Petra C.F.
 : APPLICANT: Bennett, Alice M.
 : TITLE OF INVENTION: VACCINES FOR PLAGUE
 : NUMBER OF SEQUENCES: 24
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: NIXON & VANDERHYE P.C.
 : STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
 : CITY: Arlington
 : STATE: VA
 : COUNTRY: USA
 : ZIP: 22201-4741
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/913,477
 : FILING DATE: 15-SEP-1997
 : CLASSIFICATION: 424
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/GB96/00571
 : FILING DATE: 13-MAR-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 9524825.7
 : FILING DATE: 13-MAR-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 9518946.0
 : FILING DATE: 15-SEP-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 9524825.8
 : FILING DATE: 05-DEC-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Crawford, Arthur R.
 : REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 124-599
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-816-4000
 TELEFAX: 703-816-4100
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1014 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Yersinia pestis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..987
 US-08-913-477-1

Query Match 59.2%; Score 927; DB 4; Length 1014;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

589 ATTAGAGCCTACGACAAACCCACACATTTTATGAGGATCTAGAAAAAGTTAGGCTG 648
 Db 13 ATTAGAGCCTACGACAAACCCACACATTTTATGAGGATCTAGAAAAAGTTAGGCTG 72

649 GAACAACTTACTGTCATGGTTCTTCAGTTTGAAGAATGGTTCACTTACGAAAGAT 708
 Db 73 GAACAACTTACTGTCATGGTTCTTCAGTTTGAAGAATGGTTCACTTACGAAAGAT 132

709 AAAAATATAGATTTTCCATTAAATATGATCCGAAAGATCGGAGTTTGGCAAT 768
 Db 133 AAAAATATAGATTTTCCATTAAATATGATCCGAAAGATCGGAGTTTGGCAAT 192

769 AGAGTAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828
 Db 193 AGAGTAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252

829 GATACCATTTTAAAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 888
 Db 253 GATACCATTTTAAAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 312

889 AAAGAGTTCCTTAATCATGCGGCAATGCAATGGAATTTGGGGCTTCATGGAGTA 948
 Db 313 AAAGAGTTCCTTAATCATGCGGCAATGCAATGGAATTTGGGGCTTCATGGAGTA 372

949 ATGCATTTCTCTTTAAACCGCGGATGATGATGATGATGATGATGATGATGATGAT 1008
 Db 373 ATGCATTTCTCTTTAAACCGCGGATGATGATGATGATGATGATGATGATGATGAT 432

1009 TCATGATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
 Db 433 TCATGATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492

1069 GCCGAATTAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
 Db 493 GCCGAATTAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552

1129 GGCACATAAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
 Db 553 GGCACATAAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 612

1189 ACAGATGAGAGATTTTAAAGCGGAGAGTACAAAATTTCTCGAAGAAATGCTCAA 1248
 Db 613 ACAGATGAGAGATTTTAAAGCGGAGAGTACAAAATTTCTCGAAGAAATGCTCAA 672

1249 ACCACATTCAGGTGGATGGAGCGAGAAAAAATGTCGATTAAGGACTTCTTGGG 1308
 Db 673 ACCACATTCAGGTGGATGGAGCGAGAAAAAATGTCGATTAAGGACTTCTTGGG 732

OY	1309	AGTAGAGATTAAGAACCGGGGGCTGGGTAATCTGAAAACTCAATCTATATATAA	1368
Db	733	AGTAGAGATTAAGAACCGGGGGCTGGGTAATCTGAAAACTCAATCTATATATAA	792
OY	1369	GATATATATGATATATATCTCACTTTGGCCACACCTGCTGGATTAAGTCCAGGCGCGTCAAC	1428
Db	793	GATATATATATATATATCTCACTTTGGCCACACCTGCTGGATTAAGTCCAGGCGCGTCAAC	852
OY	1429	GACTTGGTTAGCCAAAAAACAATCTAGCTGTGATATATACATCAAGTTTAAATTCAGCT	1488
Db	853	GACTTGGTTAGCCAAAAAACAATCTAGCTGTGATATATACATCAAGTTTAAATTCAGCT	912
OY	1489	ATTGACACCACTGAACCGTTTCATTTCAGAAATATGATATCAGTATGATCCAAAGTCTGCTAGAT	1548
Db	913	ATTGACACCACTGAACCGTTTCATTTCAGAAATATGATATCAGTATGATCCAAAGTCTGCTAGAT	972
OY	1549	GACACGTCGTGGTAAATGA	1566
Db	973	GACACGTCGTGGTAAATGA	990

RESULT 2
US-08-913-477-16
Sequence 16, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Peter C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
City: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ. ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1462 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```

:      TOPOLOGY: linear
:      MOLECULE TYPE: CDNA
:      HYPOTHETICAL: NO
:      ANTI-SENSE: NO
:      ORIGINAL SOURCE:
:      ORGANISM: Yersinia pestis
:      FEATURE:
:      NAME/KEY: CDS
:      LOCATION: 8..1447
:
US-08-913-477-16
```

	Query Match Similarity	59.2%	Score 927:	DB 4:	Length 1462:	
	Pred Local Similarity	99.9%	Pred. No. 0:			
	Matches 977:	Conservative 0:	Mismatches 1:	Indels 0:	Gaps 0:	
OY	ATTAGAGCCTACGAAACAACCACACATTTTATTGTAGATCTAGAAAAAGTTAGGGTG	648				
Db	ATTAGAGCCTACGAAACAACCACACATTTTATTGTAGAGATCTAGAAAAAGTTAGGGTG	532				
OY	GAAACAATTACTGTCATCGTTCCTCACTTTTAGAAGAATTGGTTCAGTTAGTCAAAGAT	708				
Db	GAAACAATTACTGTCATCGTTCCTCACTTTTAGAAGAATTGGTTCAGTTAGTCAAAGAT	592				
OY	AAAATATAGATATTTCATTTAAATATATATCCAAAAAAGATTGGAGTTTTGCCAAT	768				
Db	AAAATATAGATATTTCATTTAAATATATATCCAAAAAAGATTGGAGTTTTGCCAAT	652				
OY	AGAGTAATTACTGATGATATCGAATTGCATGAAGAAATCCCTAGTTATTCTTACCCGAG	828				
Db	AGAGTAATTACTGATGATATCGAATTGCATGAAGAAATCCCTAGTTATTCTTACCCGAG	712				
OY	GATACCATTTCTTAAAGCGGCTATTATGACAACAACCTGCAAAATGGCATCAAGCGAGTA	888				
Db	GATGCAATTTCTTAAAGCGGCTATTATGACAACAACCTGCAAAATGGCATCAAGCGAGTA	772				
OY	AAAGAGTTCCTTGATCATATCCGCCAATACACAATGGGAATTGGCGGCTCATGGCAGTA	948				
Db	AAAGAGTTCCTTGATCATATCCGCCAATACACAATGGGAATTGGCGGCTCATGGCAGTA	832				
OY	ATGCATTTCTCTTTAACCGCCGATGTATTCGATGATGATATTTTAAAGTATTTGTGAT	1008				
Db	ATGCATTTCTCTTTAACCGCCGATGTATTCGATGATGATATTTTAAAGTATTTGTGAT	892				
OY	TCAATGAATCATCATGATGATGATGCCCCGTAGCAAGTTGCGTGAAGAATTAGCTGAGCTTACC	1068				
Db	TCAATGAATCATCATGATGATGATGCCCCGTAGCAAGTTGCGTGAAGAATTAGCTGAGCTTACC	952				
OY	GCCGAATTAAGAATTTATTTCAGTTATTCAGCCGAAATTAATAGCATCTGTCTAGTAGT	1128				
Db	GCCGAATTAAGAATTTATTTCAGTTATTCAGCCGAAATTAATAGCATCTGTCTAGTAGT	1012				
OY	GGCACATTAATAATCCATGATTAATCCATTTATTCATGAGATAAAAATTATATAGTTAT	1188				
Db	GGCACATTAATAATCCATGATTAATCCATTTATTCATGAGATAAAAATTATATAGTTAT	1072				
OY	ACAGATGAAGAATTTTAAAGCCAGCCAGAGTACAAAAATCTCGAGAAAAATGCTCTCAA	1248				
Db	ACAGATGAAGAATTTTAAAGCCAGCCAGAGTACAAAAATCTCGAGAAAAATGCTCTCAA	1132				
OY	ACCACCATTCAGGTGATGATGAGGAGAGAAAAAATAGTCTCGATTAAGAAGACTTTCTTGGA	1308				
Db	ACCACCATTCAGGTGATGATGAGGAGAGAAAAAATAGTCTCGATTAAGAAGACTTTCTTGGA	1192				
OY	AGTGAATTAAGAAGAACCGGGCGTTGGGTATCTGAAAAAACTCATCTTATATATAAA	1368				
Db	AGTGAATTAAGAAGAACCGGGCGTTGGGTATCTGAAAAAACTCATCTTATATATAAA	1252				
OY	GATTAATTAATGAATTAATCTCATCTTTGCCACACACCTCTCGGATTAATGCCAGCGCTCAAC	1428				
Db	GATTAATTAATGAATTAATCTCATCTTTGCCACACACCTCTCGGATTAATGCCAGCGCTCAAC	1312				
OY	GACTTGTTAGCCAAAAACAACCTCAGCTGTCTGATTAATTACATCAAGTTTATATTACAGCT	1488				

DB 1313 GACTGTGTTAGCCAAAAACACTGCTGTGATATTACACGTTTAAATTCAGCT 1372
QY 1489 ATTGAAGCACTGAACCGTTTCATTGAGAAATATGATTCAGTATGCAAGCTGTAGAT 1548
DB 1373 ATTGAAGCACTGAACCGTTTCATTGAGAAATATGATTCAGTATGCAAGCTGTAGAT 1432
QY 1549 GACACGCTGTGTAATGA 1566
DB 1433 GACACGCTGTGTAATGA 1450

RESULT 3
US-08-913-477-22
Sequence 22, Application US/08913477
Patent No. 5985285

GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 1530 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1515

US-08-913-477-22

Query Match 59.2%; Score 927; DB 4; Length 1530;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 589 ATTGAGCCTAGACAAAACCCAAACATTTTATGAGATCTAGAAAAGTTAGGTTG 648
DB 541 ATTGAGCCTAGACAAAACCCAAACATTTTATGAGATCTAGAAAAGTTAGGTTG 600
QY 649 GAACAACTTACTGTCATGTTCTTCACTTTTGAAGAATTGGTTAGTCAAGAT 708
DB 601 GAACAACTTACTGTCATGTTCTTCACTTTTGAAGAATTGGTTAGTCAAGAT 660
QY 709 AAAAATATGATTTTCCATTAATATGATCCAGAAAAGATTGGAGTTTGGCAAT 768
DB 661 AAAAATATGATTTTCCATTAATATGATCCAGAAAAGATTGGAGTTTGGCAAT 720
QY 769 AGAGTAATTACTGATGATATCGAATTGCTCAAGAAAATCCTAGCTTATTTCTACCCGAG 828
DB 721 AGAGTAATTACTGATGATATCGAATTGCTCAAGAAAATCCTAGCTTATTTCTACCCGAG 780
QY 829 GATACATTTCTTAAGCGCGCTATTATGACAACTGCAAAAATGCGATCAAGCAGTA 888
DB 781 GATGCAATTTCTTAAGCGCGCTATTATGACAACTGCAAAAATGCGATCAAGCAGTA 840
QY 889 AAGAGTTCTTGAATCATCGCGGATGACAAATGGAATGGGGCGTTCAATGGCAGTA 948
DB 841 AAGAGTTCTTGAATCATCGCGGATGACAAATGGAATGGGGCGTTCAATGGCAGTA 900
QY 949 ATGCATTTCTTTAAGCGCGGATGATCATGATGATATTTTGAAGTGTGAT 1008
DB 901 ATGCATTTCTTTAAGCGCGGATGATCATGATGATATTTTGAAGTGTGAT 960
QY 1009 TCATATATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
DB 961 TCATATATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1069 GCGCAATTAAGATTATTCAGTATTCAAGCGGAAATTAATAGCATGCTGATGAT 1128
DB 1021 GCGCAATTAAGATTATTCAGTATTCAAGCGGAAATTAATAGCATGCTGATGAT 1080
QY 1129 GGCACCATTAATATCCATATTAATCCATTAATTCATGATGATTAATATGATGAT 1188
DB 1081 GGCACCATTAATATCCATATTAATCCATTAATTCATGATGATTAATATGATGAT 1140
QY 1189 ACAGATGAAGATTTTAAAGCCAGCGAGTACAAAATTCGAGAAAATGCTCAA 1248
DB 1141 ACAGATGAAGATTTTAAAGCCAGCGAGTACAAAATTCGAGAAAATGCTCAA 1200
QY 1249 ACCGACATTCAGGTGATGGAGCGGAAAAAATGCTCGAATAAGACATTCCTGGA 1308
DB 1201 ACCGACATTCAGGTGATGGAGCGGAAAAAATGCTCGAATAAGACATTCCTGGA 1260
QY 1309 AGTGAATTAAGAAAGACCGGGGCTGGTAAATCTGAAAACCTCATCTTATATAA 1368
DB 1261 AGTGAATTAAGAAAGACCGGGGCTGGTAAATCTGAAAACCTCATCTTATATAA 1320
QY 1369 GATAATTAAGATTAATCTCATCTTGCACACCTGCTGGATTAAGTCAAGCGGCTAAC 1428
DB 1321 GATAATTAAGATTAATCTCATCTTGCACACCTGCTGGATTAAGTCAAGCGGCTAAC 1380
QY 1429 GACTGTGTTAGCCAAAAACACTGCTGTGATATTACACGTTTAAATTCAGCT 1488
DB 1381 GACTGTGTTAGCCAAAAACACTGCTGTGATATTACACGTTTAAATTCAGCT 1440
QY 1489 ATTGAAGCACTGAACCGTTTCATTGAGAAATATGATTCAGTATGCAAGCTGTAGAT 1548
DB 1441 ATTGAAGCACTGAACCGTTTCATTGAGAAATATGATTCAGTATGCAAGCTGTAGAT 1500
QY 1549 GACACGCTGTGTAATGA 1566

Db 1501 GACACGTCTGTAAATGA 1518

RESULT 4

US-08-913-477-3

Sequence 3, Application US/08913477

Patent No. 5985285

GENERAL INFORMATION:

APPLICANT: Titball, Richard W.

APPLICANT: Williamson, Ethel D.

APPLICANT: Leary, Sophie E.C.

APPLICANT: Bennett, Alice M.

TITLE OF INVENTION: VACCINES FOR PLAGUE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHIVE P.C.

STREET: 1100 No. 5985285th Glade Rd.

CITY: Arlington

STATE: VA

COUNTRY: USA

ZIP: 22201-4741

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/913.477

FILING DATE: 15-SEP-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB96/00571

FILING DATE: 13-MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9505059.7

FILING DATE: 13-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9518946.0

FILING DATE: 15-SEP-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9524825.8

FILING DATE: 05-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Crawford, Arthur R.

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 124-599

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1014 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Yersinia pestis

FEATURE:

NAME/KEY: CDS

LOCATION: 1..987

US-08-913-477-3

Query Match

Best Local Similarity 58.9%; Score 923; DB 4; Length 1014;

Matches 973; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

593 GAGCCTAGCAAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGGTGAAC 652

Db 17 GAGCCTAGCAAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGGTGAAC 76

QY 653 AACTACGTCATGCTCTTCATGTTAGAAAGATTTGTTACGTACTCAAGATTA 712

Db 77 AACTACGTCATGCTCTTCATGTTAGAAAGATTTGTTACGTACTCAAGATTA 136

QY 713 ATATGATATTTTCCATTAATATGATCCAGAAAAGATTCGAGGTTTGGCATTAG 772

Db 137 ATATGATATTTTCCATTAATATGATCCAGAAAAGATTCGAGGTTTGGCATTAG 196

QY 773 TAATACGATGATATGCAATTTCTCAAGAAAATCCTAGTTATTTTCTACCCGAGATA 832

Db 197 TAATACGATGATATGCAATTTCTCAAGAAAATCCTAGTTATTTTCTACCCGAGATA 256

QY 833 CCATCTTAAAGCGCGTATATGCAACCAACTGCCAAAATGGCATCAAGCGATTAAG 892

Db 257 CCATCTTAAAGCGCGTATATGCAACCAACTGCCAAAATGGCATCAAGCGATTAAG 316

QY 893 AGTTCCTGAATCATCGCCGCAATATACAAATGGGAATTCGGGCGTTCAATGGCATATGC 952

Db 317 AGTTCCTGAATCATCGCCGCAATATACAAATGGGAATTCGGGCGTTCAATGGCATATGC 376

QY 953 ATTTCCTTTAACCGCCGATCGATGATGATATTTTGAAGTATTTGTTGATTC 1012

Db 377 ATTTCCTTTAACCGCCGATCGATGATGATATTTTGAAGTATTTGTTGATTC 436

QY 1013 TGAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072

Db 437 TGAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 496

QY 1073 AATTAAGATTTATTCAGTTATTCAGGCAATTAATTAAGATTTATTCAGTTATTCAG 1132

Db 497 AATTAAGATTTATTCAGTTATTCAGGCAATTAATTAAGATTTATTCAGTTATTCAG 556

QY 1133 CCATTAATATTCATGATTAATTCATGATGATGATGATGATGATGATGATGATGATGAT 1192

Db 557 CCATTAATATTCATGATTAATTCATGATGATGATGATGATGATGATGATGATGATGAT 616

QY 1193 ATGAAGATTTTAAAGCAGCGAGTACAAATTTCTGAGAAAATGCCCTAAACCA 1252

Db 617 ATGAAGATTTTAAAGCAGCGAGTACAAATTTCTGAGAAAATGCCCTAAACCA 676

QY 1253 CCATCAGGTGATGAGCGGAGGAAATTAATTAAGATTTATTCAGTTATTCAGTTATTCAG 1312

Db 677 CCATCAGGTGATGAGCGGAGGAAATTAATTAAGATTTATTCAGTTATTCAGTTATTCAG 736

QY 1313 AGAATTAAGAAACCGGGCGTTGGTATCTGAATAAATCTCATCTATATAAAGATA 1372

Db 737 AGAATTAAGAAACCGGGCGTTGGTATCTGAATAAATCTCATCTATATAAAGATA 796

QY 1373 ATATGATATTTATCTACCTTTCACACACCTGCTGGATTAAGTCAGGCGCTCAAGACT 1432

Db 797 ATATGATATTTATCTACCTTTCACACACCTGCTGGATTAAGTCAGGCGCTCAAGACT 856

QY 1433 TGGTATGCAAAACCACTAGCTGCTGATATTAATTAAGTATTAATTAAGTATTAAGTATTA 1492

Db 857 TGGTATGCAAAACCACTAGCTGCTGATATTAATTAAGTATTAATTAAGTATTAAGTATTA 916

QY 1493 AAGCAGTAAACCGTTTCAATGAAATATGATGATGATGATGATGATGATGATGATGATGAT 1552

Db 917 AAGCAGTAAACCGTTTCAATGAAATATGATGATGATGATGATGATGATGATGATGATGAT 976

QY 1553 CGCTGTGTAATGA 1566

Db 977 CGCTGTGTAATGA 990

RESULT 5

US-08-913-477-20

Sequence 20, Application US/08913477

Patent No. 5985285

GENERAL INFORMATION:

APPLICANT: Titball, Richard W.

Wed Aug 23 11:45:56 2000

us-08-699-716a-1.oliqo.rni

Page 17

```

:      TELEX: 3723836
:      INFORMATION FOR SEQ ID NO: 44.:
:      SEQUENCE CHARACTERISTICS:
:      LENGTH: 39 base pairs
:      TYPE: nucleic acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:
US-08-432-871C-44

```

```

Query Match      1.8%; Score 28; DB 3; Length 39;
Best Local Similarity 100.0%; Pred. No. 9 4e-05;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      19 CATCATCATCATCATCACACGACGCGGCC 46
          |||||||||||||||||||||||||||
DB       36 CATCATCATCATCATCACACGACGCGGCC 9

```

Search completed: August 22, 2000, 16:29:01
Job time: 3138 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 15:10:37 ; Search time 614.39 Seconds
(without alignments)
11240.077 Million cell updates/sec

Title: US-08-699-716A-1
Perfect score: 1566
Sequence: 1 ATGGCCATCATCATCATCA.....ATGACACGCTGTGTAATGA 1566

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 5247842 seqs, 2204914090 residues

Word size : 0

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_est56:*
94: gb_est57:*
95: gb_est58:*
96: gb_est59:*
97: em_est60:*
98: em_est61:*
99: em_est62:*
100: em_est63:*
101: gb_est64:*
102: gb_est65:*
103: gb_est66:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: em_est70:*
108: em_est71:*
109: em_est72:*
110: em_est73:*
111: em_est74:*
112: em_est75:*
113: gb_est76:*
114: gb_est77:*
115: em_est78:*
116: gb_est79:*


```

713 1.4 356 28 AA597472
714 1.4 358 39 AI437800
715 1.4 360 39 AI440948
716 1.4 361 39 AI441977
717 1.4 366 35 AI099882
718 1.4 366 80 AW711925
719 1.4 367 39 AI460454
720 1.4 369 40 AI495228
721 1.4 370 40 AI476864
722 1.4 373 40 AI522921
723 1.4 374 39 AI443245
724 1.4 375 39 AI441002
725 1.4 376 39 AI441959
726 1.4 376 40 AI496380
727 1.4 378 39 AI441419
728 1.4 378 86 N21757
729 1.4 381 90 W06763
730 1.4 382 81 D24300
731 1.4 383 39 AI443189
732 1.4 385 104 AO572155
733 1.4 387 93 AO078124
734 1.4 390 70 AM264812
735 1.4 391 40 AI496675
736 1.4 392 48 AU065387
737 1.4 393 63 AM098935
738 1.4 394 80 AW711094
739 1.4 395 39 AI437782
740 1.4 397 40 AI460777
741 1.4 399 80 AW710304
742 1.4 399 80 AW713441
743 1.4 399 80 AW714575
744 1.4 399 80 AW714596
745 1.4 403 40 AI494731
746 1.4 404 45 AI898615
747 1.4 405 40 AI488683
748 1.4 408 35 AI099814
749 1.4 409 39 AI437715
750 1.4 409 39 AI443076
751 1.4 409 93 AO074711
752 1.4 410 38 AI354000
753 1.4 413 39 AI442444
754 1.4 421 40 AI495640
755 1.4 422 73 AM529343
756 1.4 423 40 AI496476
757 1.4 425 40 AI495328
758 1.4 425 40 AI507820
759 1.4 427 46 AI943337
760 1.4 428 63 AM100395
761 1.4 429 63 AM052953
762 1.4 430 118 AO972390
763 1.4 433 93 AO064367
764 1.4 434 37 AI263008
765 1.4 436 43 AI748167
766 1.4 436 80 AW711210
767 1.4 437 40 AI507831
768 1.4 438 74 AM597587
769 1.4 439 44 AI779552
770 1.4 443 63 AM086922
771 1.4 441 64 AM159411
772 1.4 444 39 AI437930
773 1.4 444 74 AM585554
774 1.4 445 79 AM666050
775 1.4 445 92 Z47664
776 1.4 448 45 AI1900612
777 1.4 451 35 AI104057
778 1.4 453 29 AA675869
779 1.4 455 74 AM595952
780 1.4 457 63 AM102259
781 1.4 457 74 AM595967
782 1.4 457 89 T22782
783 1.4 458 74 AM598747
784 1.4 461 113 AO737038
785 1.4 463 41 AI575193

AA597472 29743 Lam
AI437800 SA39D09.Y
AI440948 SA55508.Y
AI441977 SA82E10.Y
AI099882 34035 Lam
AW711925 f6907ne.f
AI460454 SA78903.Y
AI495228 SB02C05.Y
AI476864 fP52A05.Y
AI522921 SA92A12.Y
AI443245 SA46C03.Y
AI441002 SA63F09.Y
AI441959 SA83C08.Y
AI496380 SB05I12.Y
AI441419 SA59F10.Y
N21757 SMNRDA0170
W06763 SMEST0399.S
D24300 RICR1688A.R
AI443189 SA45F02.Y
AO572155 HS_2108.B
AO078124 CTF-HSP-2
AM264812 X435G04.X
AI496675 SB09H11.Y
AU065387 AU065387
AM098935 GA09D06.Y
AW711094 e8G09ne.f
AI437782 SA39E06.Y
AI460777 SA69E409.Y
AW710304 e2G10ne.f
AW713441 g8H05ne.f
AW714575 l1C12ne.f
AW714596 l1E04ne.f
AI494731 SB14E07.Y
AI898615 EST268058
AI488683 EST247022
AI099814 33967 Lam
AI437715 SA83F10.Y
AI443076 SA47B01.Y
AO074711 CTF-HSP-2
AI354000 Z6H1264.S
AI442444 SA27E05.Y
AI495640 SB14I12.Y
AM529343 UT-R-BT1-
AI496476 SB08B11.Y
AI495328 SA91A11.Y
AI507820 SA87F102.Y
AI943337 fC78H05.Y
AM100395 SA27B12.Y
AM052953 634076H12
AO972390 RPCI-23-3
AO064367 HS_2205.A
AI263008 Q434E11.X
AI748167 SB49A06.Y
AW711210 e9E10ne.f
AI507831 SA87F02.Y
AM597587 349E189.Y
AI779552 EST260431
AM086922 GA09D06.X
AM159411 ZA89H04.X
AI437930 SA41E04.Y
AM585554 EST31717
AM666050 SK31E02.Y
Z47664 ATTS4520.St
AI1900612 SC13F08.Y
AI104057 EST213346
AA675869 MPAFCW5G0
AM595952 S196A03.Y
AM102259 SA85D09.Y
AM595967 EST314615
AM598747 sJ46d11.Y
AO737038 HS_2244.B
AI575193 UT-R-G0-U

```



```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

RESULT      2
LOCUS       AM154180/c
DEFINITION  F123elD.0.s1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone
              2602026 3' similar to contains element LTR1 repetitive element ; ,
              mRNA sequence.
ACCESSION   AM154180
VERSION     AM154180.1
KEYWORDS    EST.
SOURCE      Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
ORGANISM    zebrafish.
REFERENCE   1 (bases 1 to 595)
AUTHORS     Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
            Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Rucaba,T.,
            Martin,J.E., Pape,D., Steptoe,M., Underwood,K., Theising,B.,
            Ratter,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
            Mashu zebrafish EST Project 1999
JOURNAL     Unpublished (1999)
COMMENT     On Jun 5, 1998 this sequence version replaced gi:3189302.
            Other-ESTs: f123elD.y1
            Contact: S.L. Johnson
            Washington University School of Medicine
            444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watscn.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing By: Washington University Genome Sequencing Center
            Zebrafish Identity (P-value greater than 1e-99) found to :
            g112829333(g1AA542448|AA542448 fa07a10.s1 zebrafish ICRFzfs Danio
            rerio CDNA
            Seq primer: T7 ET from Amersham
            High quality sequence stop: 515.
FEATURES
             source
              Location/Qualifiers
                1..595
                 /organism="Danio rerio"
                 /strain="AB"
                 /db_xref="taxon:7955"
                 /clone="2602026"
                 /clone_id="Sugano Kawakami zebrafish DRA"
                 /sex="mixed (one male and one female, including
                   unfertilized eggs)"
                 /dev_stage="adult"
                 /lab_host="DH10B (phage resistant)"
                 /note="Vector: pME18S-FR3; Site_1: DraIII (CACTGTG);
                   Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                   with an oligo(dT) primer [ATGTGGCCCTTTTTTTTTTTTT];
                   double-stranded cDNA was ligated to a DraIII adaptor
                   [TGTTGGCCCTACTGTG], digested and cloned into distinct DraIII
                   sites of the pME18S-FR3 vector (5' site CACTGTG, 3' site
                   CACATGTG). XhoI should be used to isolate the cDNA
                   insert. Size selection was performed to exclude fragments
                   <1.5kb. Library constructed by Dr. Sumio Sugano
                   (University of Tokyo Institute of Medical Science) and
                   kindly donated by Dr. Koichi Kawakami. Custom primers for
                   sequencing: 5' end primer CTCTGCCTTAAGCTCGC and 3' end
                   primer CGACCTCGACCTGACGAC."
BASE COUNT  117 a      146 c      154 g      178 t
ORIGIN
Query Match          2.1%; Score 33; DB 64; Length 595;
Best Local Similarity 100.0%; Pred. No. 7e-07;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


JOURNAL
MEDLINE
99380589
COMMENT

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
On Feb 19, 1999 this sequence version replaced g1:411707.
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.hsc.washington.edu>
Plate: 3135 Row: E Column: 20
Seq primer: T7
Class: BAC ends
High quality sequence stop: 744.

FEATURES
Source

1..744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-3135 Col-20 Row-E"
/clone_id="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC clones in E-Coli DH10B"
BASE COUNT 224 a 102 c 152 g 265 t 1 others
ORIGIN

Query Match 2.0%; Score 32; DB 117; Length 744;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 CCATCATCATCATCATCATCATCATCATCATCA 37
|||||
Db 464 CCATCATCATCATCATCATCATCATCATCA 433

RESULT 6
A0539857/c 245 bp DNA GSS 19-MAY-1999
LOCUS
DEFINITION RPCR-11-31614.TV RPCR-11 Homo sapiens genomic clone RPCR-11-31614, genomic survey sequence.
ACCESSION A0539857

VERSION A0539857.1 GI:4870387
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC end sequences from library RPCR-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other GSS: RPCR-11-31614.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeef@igir.org

Clones are derived from the human BAC library RPCR-11. For BAC library availability, please contact Pieter de Jong (pieter@ejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7.
Class: BAC ends.

FEATURES
Source

Location/Qualifiers
1..245
/organism="Homo sapiens"
/db_xref="GDB:762115"
/db_xref="taxon:9606"
/clone="RPCR-11-31614"
/clone_id="RPCR-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACs.6; Site_1: EcoRI; Site_2: EcoRI; RPCR11 Human Male BAC library"
BASE COUNT 70 a 40 c 49 g 86 t
ORIGIN

Query Match 2.0%; Score 31; DB 104; Length 245;
Best Local Similarity 100.0%; Pred. No. 8.3e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 CATCATCATCATCATCATCATCATCATCATCA 37
|||||
Db 243 CATCATCATCATCATCATCATCATCATCA 213

RESULT 7
AM624044/c 424 bp mRNA EST 28-MAR-2000
LOCUS
DEFINITION EST321989 tomato flower buds 3-8 mm, Cornell University
Lycopersicon esculentum cDNA clone CTOB13N18 5', mRNA sequence.
ACCESSION AM624044
VERSION AM624044.1 GI:7337071
KEYWORDS EST.
SOURCE tomato.

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 424)
van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE Generation of ESTs from tomato flower tissue, 3-8 mm buds
JOURNAL Unpublished (1999)
COMMENT On Apr 30, 1999 this sequence version replaced g1:4727518.

CONTACT: David Fritsch
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfritsch@CLEMSON.EDU

FEATURES
Source

Location/Qualifiers
1..424
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CTOB13N18"
/clone_id="tomato flower buds 3-8 mm, Cornell University"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Tanksley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."
BASE COUNT 154 a 73 c 96 g 101 t
ORIGIN

Query Match 2.0%; Score 31; DB 79; Length 424;
Best Local Similarity 100.0%; Pred. No. 8.6e-06;

FEATURES
SOURCE

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence

Location/Qualifiers
1. .600
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOB13N16"
/clone_id="tomato flower buds 3-8 mm, Cornell University"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Tansley; flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT
ORIGIN

203 a 100 c 132 g 164 t 1 others

Query Match 2.0%; Score 31; DB 79; Length 600;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0y 6 CCATCATCATCATCATCATCATCATCATCAC 36
|||||
Db 378 CCATCATCATCATCATCATCATCATCATCAC 348

RESULT 12
LOCUS AM649317 628 bp mRNA EST 04-APR-2000
DEFINITION EST327771 tomato germinating seedlings, TAMU Lycopersicon
esculentum cDNA clone cLEI7J2 5', mRNA sequence.
ACCESSION AM649317
VERSION AM649317.1 GI:7410555
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
Solanales; Solanum; Lycopersicon.
1 (bases 1 to 628)
Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,
Nierman, W., Fraser, C.M., Giovannoni, J.J., Martin, G.B. and
Tansley, S.D.
Generation of ESTs from germinating tomato seed
unpublished (2000)
On Jan 6, 2000 this sequence version replaced gi:6675154.
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence.

FEATURES
SOURCE

Location/Qualifiers
1. .628
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEI7J2"
/clone_id="tomato germinating seedlings, TAMU"
/tissue_type="whole seedlings"
/dev_stage="7 days post imbibition"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

Whole, 7 days post imbibition on water-agar. Mixed stage whole germinating seedlings from seed coat emergence up to two centimeters in seeds not showing obvious signs of germination were discarded."

BASE COUNT 209 a 115 c 142 g 161 t 1 others

Query Match 2.0%; Score 31; DB 79; Length 628;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CCATCATCATCATCATCATCATCATCATCATC 36
DB 518 CCATCATCATCATCATCATCATCATCATC 488

RESULT 13

C24646 650 bp mRNA EST 28-APR-1999
LOCUS C24646 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
DEFINITION discoidium cDNA clone SL-X046, mRNA sequence.
ACCESSION C24646 GI:2243067
VERSION C24646.1
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum.
REFERENCE Dictyostelium; Dictyostelida; Dictyostelium.
AUTHORS Mori, T., Urushihara, H., Salto, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Mita, B. N., Pl. M., Salto, T., Takemoto, K., Yashikawa, H., Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
TITLE The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development

JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99136227
COMMENT On Jul 9, 1999 this sequence version replaced g1:5435108.

CONTACT: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Tennodai, Tsukuba, Ibaraki 305, Japan
Email: d40zhuesakura.cc.tsukuba.ac.jp
PROJECT - 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES Location/Qualifiers
1..650
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SL-X046"
/clone_1lb="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"

BASE COUNT 231 a 92 c 124 g 203 t

Query Match 2.0%; Score 31; DB 80; Length 650;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CATCATCATCATCATCATCATCATCATC 37
DB 40 CATCATCATCATCATCATCATCATCATC 70

RESULT 14

LOCUS A0349263 113 bp DNA GSS 07-MAY-1999
DEFINITION RPCI11-139F19, TV RPCI-11 Homo sapiens genomic clone RPCI-11-139F19, genomic survey sequence.
ACCESSION A0349263

VERSION A0349263.1 GI:4174159
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 113)
Zhang, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@tigr.org, med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: 17
Class: BAC ends.

FEATURES Location/Qualifiers
1..113
/organism="Homo sapiens"
/db_xref="GDB:755130"
/db_xref="taxon:9606"
/clone="RPCI-11-139F19"
/clone_1lb="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI;
RPCI11 Human Male BAC Library"

BASE COUNT 33 a 23 c 17 g 40 t

Query Match 1.9%; Score 30; DB 101; Length 113;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ATCATCATCATCATCATCATCATCATCATC 37
DB 9 ATCATCATCATCATCATCATCATCATCATC 38

RESULT 15
LOCUS HS021500 198 bp DNA GSS 10-MAR-1997
DEFINITION Human chromosome 17q21 clone ET10, genomic survey sequence.
ACCESSION U21500
VERSION U21500.1 GI:1872174
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 198)
McFarlane, R., de Jong, P., Quirk, V.P., Lehrach, H. and Solomon, E.
Physical mapping, cloning, and identification of genes within a
500-kb region containing BRCA1
Proc. Natl. Acad. Sci. U.S.A. 92 (10), 4362-4366 (1995)
JOURNAL 95273363
MEDLINE 95273363
REFERENCE 2 (bases 1 to 198)
Brown, M.A.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1995) Melissa A. Brown, Somatic Cell Genetics,

Imperial Cancer Research Fund, Lincoln's Inn Fields, London, WC2A
3PX, UK
COMMENT On Mar 10, 1997 this sequence version replaced gi:852476.
FEATURES Location/Qualifiers
Source 1..198

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ET10"
/chromosome="17"
/map="17q21"
/note="exon-trapped clone isolated using the procedure of
Buckler et al. PNAS (1991) 88:4005-4009"
BASE COUNT 54 a 28 c 58 g 58 t
ORIGIN

Query Match 1.9%: Score 30; DB 123; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CCATCATCATCATCATCATCATCA 35
|||||
DB 129 CCATCATCATCATCATCATCATCA 100

Search completed: August 22, 2000, 16:04:40
Job time: 3243 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:02:20 ; Search time 29.04 Seconds
(Without alignments)
424.947 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGNHHHHHHSSGHIDD.....RFLQKDYDVNQRLIDPTSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 188963 seqs, 23686106 residues

Word size: 0

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database: A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243	46.6	329	1 R79961	Partial LcrV (V an
2	243	46.6	329	1 W01040	Y. pestis V antige
3	243	46.6	329	1 W01041	Y. pestis V antige
4	207	39.7	480	1 W01045	Y. pestis Fl/V ant
5	170	32.6	170	1 R76528	Yersinia pestis ca
6	170	32.6	170	1 W01043	Y. pestis Fl antlg
7	170	32.6	170	1 W59782	Amino acid sequenc
8	170	32.6	501	1 W01044	Y. pestis Fl/V ant
9	170	32.6	329	1 R79962	Partial LcrV (V an
10	167	33.1	329	1 W59788	Nucleotide sequenc
11	149	28.6	149	1 W59787	Amino acid sequenc
12	149	28.6	150	1 W59787	Amino acid sequenc
13	134	29.7	192	1 W59785	Yersinia pestis ca
14	128	24.6	151	1 R76526	Yersinia pestis ca
15	128	24.6	151	1 R76527	Y. pestis Fl antlg
16	128	24.6	151	1 W01042	Amino acid sequenc
17	113	21.7	171	1 W59786	Y. pestis Fl antlg
18	24	4.6	24	1 W19779	Amino acid sequenc
19	17	3.3	33	1 W36986	Polyhistidine-ente
20	17	3.3	54	1 W36990	HFR1 peptide. Meth
21	17	3.3	323	1 W71185	FP505 protein cont
22	17	3.3	22	1 W68098	Chlamydomonas rein
23	17	3.3	448	1 W68395	Clostridium botuli
24	17	3.3	451	1 W68395	Clostridium botuli
25	17	3.3	451	1 W68398	Clostridium botuli
26	17	3.3	452	1 W68398	Clostridium botuli
27	17	3.3	462	1 R59009	Type A neurotoxin
28	17	3.3	462	1 W68397	Clostridium botuli
29	17	3.3	462	1 W68397	Clostridium botuli
30	17	3.3	472	1 W68394	Clostridium botuli
31	17	3.3	473	1 W68393	Clostridium botuli
32	17	3.3	473	1 W68400	Clostridium botuli
33	16	3.1	21	1 W23647	Recombinant squitr

34	16	3.1	21	1 W23650	Recombinant squitr
35	16	3.1	21	1 W23653	Recombinant squitr
36	16	3.1	254	1 W37699	Protein sequence o
37	12	2.3	17	1 W3939	t2 peptide, synthe
38	11	2.1	302	1 R82097	Human Gax protein.
39	11	2.1	303	1 R82096	Rat Gax protein. R
40	11	2.1	414	1 R32020	Sequence of a euka
41	11	2.1	414	1 W65406	Human transcrip
42	10	1.9	96	1 W99830	HIV Var protein se
43	10	1.9	96	1 W99831	HIV E21,24P protei
44	10	1.9	96	1 W99832	HIV A30S protein s
45	10	1.9	96	1 W99833	HIV A30L protein s
46	10	1.9	96	1 W99834	HIV A59P protein s
47	10	1.9	96	1 W99835	HIV L68S protein s
48	10	1.9	96	1 W99836	HIV H71C protein s
49	10	1.9	96	1 W99837	HIV G75A protein s
50	10	1.9	96	1 W99838	HIV C76S protein s
51	10	1.9	335	1 W94462	Human wild-type Ho
52	10	1.9	351	1 R24393	Sequence of Histid
53	10	1.9	449	1 W56703	Human GARA-6 trans
54	10	1.9	480	1 R44551	Brain factor-1. Is
55	10	1.9	626	1 W16398	Human neuron-deriv
56	10	1.9	628	1 R32057	Apoptotic cerebral
57	10	1.9	672	1 R31216	Penicillin binding
58	10	1.9	763	1 W41734	Human TRAF-2 kinas
59	10	1.9	786	1 R47066	Sequence of Crypto
60	10	1.9	2262	1 W56737	Calcium ion channe
61	10	1.9	2266	1 R11008	Human neuronal cal
62	10	1.9	2510	1 R11007	Human neuronal cal
63	9	1.7	18	1 W73374	Rec2 protein leade
64	9	1.7	19	1 W00339	Leader sequence fo
65	9	1.7	20	1 W8594	Hexa-His/thrombin
66	9	1.7	21	1 R8745	Leader sequence fo
67	9	1.7	21	1 W08590	Fusion peptide lea
68	9	1.7	21	1 W97132	Synthetic 65K-gut
69	9	1.7	22	1 W35362	peptide expressed
70	9	1.7	23	1 W65551	His-tag leader seq
71	9	1.7	25	1 R92906	PER-15b expression
72	9	1.7	43	1 R82733	Rat CAR1 fusion p
73	9	1.7	110	1 W78908	hnRNP U protein 69
74	9	1.7	137	1 R55693	Hepes simplex vir
75	9	1.7	144	1 W47195	Human obesity prot
76	9	1.7	166	1 W00525	Human obesity prot
77	9	1.7	166	1 W00529	Human obesity prot
78	9	1.7	166	1 W00533	Human obesity prot
79	9	1.7	166	1 W00524	Human obesity prot
80	9	1.7	167	1 R92730	Murine obesity prot
81	9	1.7	167	1 W00521	Murine obesity prot
82	9	1.7	167	1 R92726	Murine obesity prot
83	9	1.7	167	1 W00517	Human obesity prot
84	9	1.7	180	1 W47081	Salmonella Sef1 p
85	9	1.7	359	1 W17791	Maize nuclear-loca
86	9	1.7	359	1 W17794	KN1 alanine scann
87	9	1.7	359	1 W17795	KN1 alanine scann
88	9	1.7	359	1 W17796	KN1 alanine scann
89	9	1.7	359	1 W17797	KN1 alanine scann
90	9	1.7	359	1 W17798	KN1 alanine scann
91	9	1.7	359	1 W17799	KN1 alanine scann
92	9	1.7	371	1 W19600	KN1 alanine scann
93	9	1.7	371	1 W73369	Epitope tagged TBP
94	9	1.7	398	1 W71071	Multiple sclerosis
95	9	1.7	398	1 W99557	Protein encoded by
96	9	1.7	402	1 R24392	Sequence of the Hi
97	9	1.7	408	1 W84315	trxA-rabbit tissue
98	9	1.7	421	1 W68541	Amino acid sequenc
99	9	1.7	423	1 W96263	Bn-3a polypeptide
100	9	1.7	423	1 W96262	topoisomerase I ex
101	9	1.7	711	1 W57321	Protein allergen I
102	8	1.5	10	1 W7231	N-terminal peptide
103	8	1.5	23	1 W15248	Nickel binding pro
104	8	1.5	60	1 R90286	H. pylori cytoplas
105	8	1.5	60	1 W20250	H. pylori cytoplas
106	8	1.5	60	1 W24617	H. pylori cytoplas

107	8	1.5	64	1	W27485	Sequence used in d
108	8	1.5	78	1	W20756	H. pylori cytoplas
109	8	1.5	96	1	R36548	Recombinant Y2X. R
110	8	1.5	221	1	W23460	Recombinant MAM su
111	8	1.5	335	1	W94465	Human HoxA1 varian
112	8	1.5	403	1	W72943	Mycobacterium tube
113	8	1.5	404	1	W72942	Mycobacterium tube
114	8	1.5	494	1	W23066	Rabbit beta-2 inte
115	8	1.5	494	1	W65108	Rabbit beta-2 inte
116	8	1.5	494	1	W72840	Rabbit alpha-d clo
117	8	1.5	494	1	W73350	Rabbit alpha-d clo
118	8	1.5	531	1	W47002	Glutathione-S-tran
119	8	1.5	536	1	R88597	Fork insertion mut
120	8	1.5	600	1	R88598	Fork insertion mut
121	8	1.5	802	1	W61476	Mycobacterium anti
122	8	1.5	802	1	W81746	M. tuberculosis fu
123	7	1.3	9	1	W06142	Variant adenovirus
124	7	1.3	13	1	R72792	Hexa-histidine lea
125	7	1.3	17	1	R68877	pGEX-7Hpro proteas
126	7	1.3	64	1	Y01320	Human secreted pro
127	7	1.3	73	1	Y01331	Modified K11 RNA p
128	7	1.3	73	1	Y01295	K11 RNA polymerase
129	7	1.3	73	1	Y01362	Modified K11 RNA p
130	7	1.3	84	1	W00945	CMV500-4heptadREB
131	7	1.3	91	1	W46607	Human brain protei
132	7	1.3	96	1	W38487	Streptococcus pneu
133	7	1.3	97	1	W00949	CMV500-4heptadfos
134	7	1.3	105	1	W00947	CMV500-FosB2IP(MO)
135	7	1.3	111	1	W59030	Carbocacteriocin B
136	7	1.3	111	1	W94888	CMV400-4undzip leu
137	7	1.3	113	1	W00950	Carbocacteriocin B
138	7	1.3	124	1	W78909	Rat CAR2 fusion p
139	7	1.3	129	1	W54310	Plasmid DNA encodi
140	7	1.3	129	1	W54285	Human modified cyt
141	7	1.3	135	1	W59053	Hs rF6P1 protein
142	7	1.3	135	1	W71694	R21-2Hpro-28 fusio
143	7	1.3	136	1	R85066	Wilms' tumour anti
144	7	1.3	210	1	W22883	Wilms' tumour prot
145	7	1.3	220	1	Y02634	Prot.D1/3-E7-mut(C
146	7	1.3	220	1	Y02631	Prothb126-E7-His
147	7	1.3	227	1	Y02638	Prot.D1/3-E7-His/H
148	7	1.3	227	1	Y02640	Prot.D1/3-E7-mut(C
149	7	1.3	239	1	Y02636	CLYRA-E7-His prote
150	7	1.3	249	1	W06148	Synthetic adenovir
151	7	1.3	253	1	Y02632	Prot.D1/3-E6-His/H
152	7	1.3	278	1	Y02641	Prot.D1/3-E6-His/H
153	7	1.3	282	1	Y02635	CLYRA-E6-His prote
154	7	1.3	302	1	Y05097	HIV Tat protein. p
155	7	1.3	349	1	W54311	Plectestrin homology
156	7	1.3	371	1	Y02633	Prot.D1/3-E6-E7-Hi
157	7	1.3	383	1	Y02642	Prot.D1/3-E6-E7-Hi
158	7	1.3	390	1	Y02637	CLYRA-E6E7-His pro
159	7	1.3	391	1	W23534	CDK inhibitory fus
160	7	1.3	391	1	W95094	Human p27-p16 fusi
161	7	1.3	400	1	R75203	Tyrosine phosphata
162	7	1.3	433	1	W17783	FIV integrase-lexa
163	7	1.3	433	1	W17783	FIV integrase-lexa
164	7	1.3	515	1	W55073	Streptococcus pneu
165	7	1.3	585	1	R88595	Fork insertion mut
166	7	1.3	589	1	R88596	Fork insertion mut
167	7	1.3	685	1	W88432	Disease associated
168	7	1.3	685	1	Y00915	Human serum induci
169	7	1.3	713	1	W21960	Recombinant furin
170	7	1.3	806	1	R65495	Marek's disease vi
171	7	1.3	1262	1	W13505	B. bronchiseptica
172	7	1.3	1263	1	W13503	B. pertussis adeny
173	7	1.3	1445	1	R12108	N-terminal deleted
174	7	1.3	1522	1	P93357	Sequence of the ca
175	7	1.3	1644	1	W13504	B. bronchiseptica
176	7	1.3	1645	1	W13502	B. pertussis adeny
177	7	1.3	1705	1	P94365	Sequence of part o
178	7	1.3	1706	1	R08031	Adenyl cyclase fro
179	7	1.3	1891	1	W52844	Amycolatopsis medi
180	7	1.3	3413	1	W52849	A. mediterranei ri
181	7	1.3	3729	1	W22603	Tyactone synthase
182	6	1.2	6	1	W19161	Isoelectric point
183	6	1.2	6	1	W18225	Purification tag o
184	6	1.2	6	1	W44011	Poly-histidine pep
185	6	1.2	6	1	W69961	Poly-His tag for c
186	6	1.2	6	1	W68297	Poly-His tag for c
187	6	1.2	6	1	W63024	Hexa-histidine tag
188	6	1.2	6	1	W84203	Peptide comprising
189	6	1.2	7	1	P80401	Affinity peptide c
190	6	1.2	7	1	Y03173	Human cardiac trop
191	6	1.2	8	1	R77412	Extension peptide
192	6	1.2	8	1	R95439	Polyhistidine tail
193	6	1.2	8	1	W20036	C-terminal metal b
194	6	1.2	9	1	W61479	A. fumigatus aller
195	6	1.2	8	1	R97378	Penta-histidine pe
196	6	1.2	9	1	W60080	Homo sapiens RH ep
197	6	1.2	9	1	W90199	B7.2-GHis tag fusi
198	6	1.2	10	1	R27517	Effector cell prot
199	6	1.2	10	1	R39733	First type QE-C-t
200	6	1.2	10	1	R73684	Metal-affinity dec
201	6	1.2	10	1	R74672	Polyhistidine moti
202	6	1.2	10	1	R77469	I-domain N-termina
203	6	1.2	10	1	W31496	Human DnaE II pro
204	6	1.2	10	1	W56089	Human monocycle che
205	6	1.2	10	1	W76939	Fusion immunoglobu
206	6	1.2	10	1	W82996	Human fibronectin
207	6	1.2	10	1	W89262	N-terminal hexahis
208	6	1.2	10	1	W94259	His-tag sequence.
209	6	1.2	10	1	Y04360	Salmonella flagell
210	6	1.2	11	1	R77410	Extension peptide
211	6	1.2	11	1	W53171	C-terminal peptide
212	6	1.2	11	1	W61480	A. fumigatus aller
213	6	1.2	11	1	W70581	Amino acid sequenc
214	6	1.2	11	1	W80426	Linker used in the
215	6	1.2	12	1	R07064	Transcript of plas
216	6	1.2	12	1	R45783	Farnesyltransferas
217	6	1.2	12	1	W61549	Endoprotease Xa
218	6	1.2	12	1	W82993	Human fibronectin
219	6	1.2	13	1	W61548	Endoprotease Xa
220	6	1.2	13	1	R60018	Can fi peptide wit
221	6	1.2	14	1	W39836	Light chain CDRI o
222	6	1.2	14	1	W62982	Peptide inhibitor
223	6	1.2	15	1	R60513	Hexahistidine-cont
224	6	1.2	15	1	W72895	Mycobacterium tube
225	6	1.2	16	1	R69813	Cleavage site reco
226	6	1.2	16	1	W07945	gp120 peptide 17.
227	6	1.2	16	1	W30929	Birch pollen aller
228	6	1.2	17	1	P92081	Peptide coded for
229	6	1.2	17	1	R37708	Delta14 Ser17 hCNT
230	6	1.2	17	1	R97389	CC49 VH-spacer-PLA
231	6	1.2	17	1	W95000	N-terminal histidi
232	6	1.2	18	1	R60253	Vector-encoded NF-
233	6	1.2	18	1	R69814	Site recognised En
234	6	1.2	18	1	R78152	Synthetic leader t
235	6	1.2	18	1	R95870	CR2 receptor ligand
236	6	1.2	18	1	W71424	Peptide having an
237	6	1.2	19	1	W37707	C-terminal additio
238	6	1.2	19	1	W98067	Streptococcus pyog
239	6	1.2	20	1	R73014	N-terminal peptide
240	6	1.2	20	1	W00869	T-cell receptor al
241	6	1.2	20	1	W01445	Bioelastic polymer
242	6	1.2	20	1	W76938	Fusion immunoglobu
243	6	1.2	21	1	R07063	Transcript of plas
244	6	1.2	21	1	P71703	Internal fragment
245	6	1.2	21	1	R60142	pGEX-CHpro proteas
246	6	1.2	21	1	W48442	Myc-his peptide ta
247	6	1.2	21	1	W54281	Myc-histidine pep
248	6	1.2	21	1	W54435	Human PS112 expres
249	6	1.2	21	1	W75875	Myc-his tag peptid
250	6	1.2	21	1	W79683	Human CS198 protei
251	6	1.2	21	1	W86030	Myc-His peptide ta
252	6	1.2	21	1	W95563	Myc-his tag used i

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:03:20 ; Search time 29.04 Seconds

(without alignments)
424.947 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGNHHHHHHSSGHIDDD.....RFIQKDYDVMQRLDPTSGK 521

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 188963 seqs, 23686106 residues

Word size: 0

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database: A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	46.6	329	1 R79961	Partial LcrV (V an
2	243	46.6	329	1 W01040	Y. pestis V antige
3	243	46.6	329	1 W01041	Y. pestis V antige
4	207	39.7	480	1 W01045	Y. pestis Fl/V ant
5	170	32.6	170	1 R76528	Yersinia pestis ca
6	170	32.6	170	1 W01043	Y. pestis Fl antig
7	170	32.6	170	1 W59782	Amino acid sequenc
8	170	32.6	170	1 W59783	Amino acid sequenc
9	170	32.6	501	1 W01044	Y. pestis Fl/V ant
10	167	32.1	329	1 R79962	Partial LcrV (V an
11	149	28.6	149	1 W59788	Nucleotide sequenc
12	149	28.6	150	1 W59787	Amino acid sequenc
13	134	22.7	192	1 W59785	Yersinia pestis ca
14	128	24.6	151	1 R76526	Yersinia pestis ca
15	128	24.6	151	1 R76527	Y. pestis Fl antig
16	128	24.6	151	1 W01042	Amino acid sequenc
17	113	21.7	171	1 W59786	Y. pestis Fl antig
18	24	4.6	24	1 W18779	Amino acid sequenc
19	17	3.3	53	1 W36986	Polyhistidine-ente
20	17	3.3	54	1 W36990	HFR peptide. Meth
21	17	3.3	323	1 W71185	FP505 protein cont
22	17	3.3	423	1 W68098	Chlamydomonas rein
23	17	3.3	448	1 W68399	Clostridium botuli
24	17	3.3	451	1 W68399	Clostridium botuli
25	17	3.3	451	1 W68399	Clostridium botuli
26	17	3.3	452	1 W68396	Clostridium botuli
27	17	3.3	462	1 R95009	Clostridium botuli
28	17	3.3	462	1 W68390	Type A neurotoxin
29	17	3.3	462	1 W68397	Clostridium botuli
30	17	3.3	472	1 W68397	Clostridium botuli
31	17	3.3	472	1 W68397	Clostridium botuli
32	17	3.3	473	1 W68400	Clostridium botuli
33	16	3.1	21	1 W23647	Recombinant squitr

34	16	3.1	21	1 W23650	Recombinant squitr
35	16	3.1	21	1 W23653	Recombinant squitr
36	16	3.1	254	1 W37699	Protein sequence o
37	12	2.3	17	1 W03939	T2 peptide, synthe
38	11	2.1	302	1 R82097	Human Gax protein.
39	11	2.1	303	1 R82096	Rat Gax protein. R
40	11	2.1	414	1 R32020	Sequence of a euka
41	11	2.1	414	1 W65406	Human transcriptio
42	10	1.9	96	1 W99830	HIV Vpr protein se
43	10	1.9	96	1 W99831	HIV E21, 2AP protei
44	10	1.9	96	1 W99832	HIV A30S protein s
45	10	1.9	96	1 W99833	HIV A30L protein s
46	10	1.9	96	1 W99834	HIV A59P protein s
47	10	1.9	96	1 W99835	HIV L68S protein s
48	10	1.9	96	1 W99836	HIV C76S protein s
49	10	1.9	96	1 W99837	HIV C76S protein s
50	10	1.9	96	1 W99838	Human wild-type Ho
51	10	1.9	335	1 W94462	Sequence of Histid
52	10	1.9	351	1 R24393	Human GATR-6 trans
53	10	1.9	449	1 W56703	Brain factor-1. IS
54	10	1.9	480	1 R44551	Human neuron-deriv
55	10	1.9	626	1 W16398	Apoptotic cerebral
56	10	1.9	628	1 R32057	Penicillin binding
57	10	1.9	672	1 R31216	Human TRAF-2 kinas
58	10	1.9	763	1 W41734	Sequence of Crypto
59	10	1.9	786	1 R47066	Calcium ion channe
60	10	1.9	2262	1 W56737	Human neuronal cal
61	10	1.9	2266	1 R11008	Human neuronal cal
62	10	1.9	2510	1 R11007	Human neuronal cal
63	9	1.7	18	1 W73374	Rec2 protein leade
64	9	1.7	19	1 W00339	Leader sequence fo
65	9	1.7	20	1 W08594	Hexa-His/thrombin
66	9	1.7	21	1 R87745	Leader sequence fo
67	9	1.7	21	1 W08590	Fusion peptide lea
68	9	1.7	21	1 W97132	Synthetic 65K-giut
69	9	1.7	22	1 W35362	Peptide expressed
70	9	1.7	23	1 W65551	His-Tag leader seq
71	9	1.7	25	1 R92906	per-1b expression
72	9	1.7	43	1 R92733	Rat CART1 fusion p
73	9	1.7	110	1 W78908	Heppes simplex vir
74	9	1.7	137	1 R56993	Human obesity prot
75	9	1.7	144	1 W47195	Human obesity prot
76	9	1.7	166	1 W00525	Human obesity prot
77	9	1.7	166	1 W00528	Human obesity prot
78	9	1.7	166	1 W00533	Human obesity prot
79	9	1.7	166	1 W00524	Human obesity prot
80	9	1.7	167	1 R92730	Human obesity prot
81	9	1.7	167	1 W00521	Human obesity prot
82	9	1.7	167	1 R92726	Human obesity prot
83	9	1.7	167	1 W00517	Human obesity prot
84	9	1.7	180	1 W47081	Salmonella Sef1a p
85	9	1.7	359	1 W17791	Maize nuclear-loca
86	9	1.7	359	1 W17794	KN1 alanine scann
87	9	1.7	359	1 W17795	KN1 alanine scann
88	9	1.7	359	1 W17796	KN1 alanine scann
89	9	1.7	359	1 W17797	KN1 alanine scann
90	9	1.7	359	1 W17798	KN1 alanine scann
91	9	1.7	359	1 W17799	KN1 alanine scann
92	9	1.7	371	1 W19600	KN1 alanine scann
93	9	1.7	371	1 W73369	Multiple sclerosis
94	9	1.7	398	1 W71071	Protein encoded by
95	9	1.7	402	1 W99557	Sequence of the H1
96	9	1.7	420	1 R24392	Trxa-rabbit tissue
97	9	1.7	420	1 W68541	Amino acid sequenc
98	9	1.7	421	1 W68541	Bm-3a polypeptide
99	9	1.7	423	1 W96263	Topoisomerase I ex
100	9	1.7	711	1 W57321	Protein allergen 1
101	9	1.7	711	1 W57321	N-terminal peptide
102	8	1.5	10	1 W23331	Nickel binding pro
103	8	1.5	23	1 W15248	H. pylori cytoplas
104	8	1.5	60	1 R90286	H. pylori cytoplas
105	8	1.5	60	1 W20250	H. pylori cytoplas
106	8	1.5	60	1 W24617	H. pylori cytoplas

107	8	1.5	64	1	W27485	Sequence used in d
108	8	1.5	78	1	W20756	R. ptyori cytoplasmic
109	8	1.5	96	1	R36548	Recombinant YZX. R
110	8	1.5	221	1	W23460	Recombinant MAM su
111	8	1.5	335	1	W94465	Human HoxA1 varian
112	8	1.5	403	1	W72943	Mycobacterium tube
113	8	1.5	404	1	W72942	Mycobacterium tube
114	8	1.5	494	1	W23066	Rabbit beta 2 inte
115	8	1.5	494	1	W65108	Rabbit beta-integr
116	8	1.5	494	1	W72840	Rabbit alpha-d clo
117	8	1.5	531	1	W73350	Rabbit alphas prot
118	8	1.5	531	1	W47002	Glutathione-S-tran
119	8	1.5	596	1	R88597	Forki insertion mut
120	8	1.5	600	1	R88598	Forki insertion mut
121	8	1.5	802	1	W64379	Mycobacterium anti
122	8	1.5	802	1	W81746	M. tuberculosis fu
123	7	1.3	9	1	W06142	Variant adenovirus
124	7	1.3	13	1	R72792	Hexa-histidine lea
125	7	1.3	17	1	R68977	pEX-7HPro proteas
126	7	1.3	64	1	Y00320	Human secreted pro
127	7	1.3	73	1	Y01331	Modified K11 RNA p
128	7	1.3	73	1	Y01295	K11 RNA polymerase
129	7	1.3	73	1	Y01362	Modified k11 RNA p
130	7	1.3	84	1	W00945	CMV500-4heptadec
131	7	1.3	91	1	W46607	Human brain protei
132	7	1.3	96	1	W38487	Streptococcus pneu
133	7	1.3	97	1	W00949	CMV500-4heptadec
134	7	1.3	105	1	W00949	CMV500-FosB2P(MO)
135	7	1.3	111	1	W59030	Carnobacteriocin B
136	7	1.3	111	1	W94888	Carnobacteriocin B
137	7	1.3	113	1	W00950	CMV400-4undecip leu
138	7	1.3	124	1	W78909	Rat CAR2 fusion p
139	7	1.3	139	1	W54310	Plasmid DNA encodi
140	7	1.3	139	1	W54285	Human modified cyt
141	7	1.3	135	1	W59053	His FEGF10 protein
142	7	1.3	136	1	W71694	R21-2Hpro-28 fusio
143	7	1.3	210	1	R85066	Wilms' tumour anti
144	7	1.3	210	1	W22883	Wilms' tumour prot
145	7	1.3	220	1	Y02634	Prot.D1/3-E7-mut(C
146	7	1.3	220	1	Y02634	Prot.D1/3-E7-mut(C
147	7	1.3	227	1	Y02638	Prot.D1/3-E7-His/H
148	7	1.3	227	1	Y02640	Prot.D1/3-E7-mut(C
149	7	1.3	239	1	Y02636	CLYTA-E7-His prote
150	7	1.3	239	1	W06148	Synthetic adenovir
151	7	1.3	273	1	Y02632	Prot.D1/3-E6-His/H
152	7	1.3	278	1	Y02641	Prot.D1/3-E6-His/H
153	7	1.3	292	1	Y02635	CLYTA-E6-His prote
154	7	1.3	302	1	Y05097	HIV Tat protein. p
155	7	1.3	349	1	W54311	Plectestrin homolo
156	7	1.3	371	1	Y02633	Prot.D1/3-E6-E7-H1
157	7	1.3	383	1	Y02642	Prot.D1/3-E6-E7-H1
158	7	1.3	390	1	Y02637	CLYTA-E6E7-His pro
159	7	1.3	391	1	W23534	CDK inhibitory fus
160	7	1.3	391	1	W95084	Human p27-p16 fusi
161	7	1.3	400	1	R75203	Tyrosine phosphat
162	7	1.3	453	1	W17784	FIV integrase-LexA
163	7	1.3	499	1	W17783	FIV integrase-LexA
164	7	1.3	515	1	W55073	Streptococcus pneu
165	7	1.3	585	1	R88595	Forki insertion mut
166	7	1.3	685	1	R88596	Forki insertion mut
167	7	1.3	685	1	W88432	Disease associated
168	7	1.3	685	1	Y00915	Human serum induc
169	7	1.3	713	1	W21960	Recombinant furin
170	7	1.3	806	1	R65495	Marek's disease vl
171	7	1.3	1262	1	W13505	B. bronchiseptica
172	7	1.3	1263	1	W13503	B. pertussis adeny
173	7	1.3	1445	1	R12108	N-terminal deleted
174	7	1.3	1522	1	P93357	Sequence of the ca
175	7	1.3	1644	1	W13504	B. bronchiseptica
176	7	1.3	1645	1	W13502	B. pertussis adeny
177	7	1.3	1705	1	P94365	Sequence of part o
178	7	1.3	1706	1	R08031	Adeny cyclase fro
179	7	1.3	1891	1	W52844	Amycolatopsis medi
180	7	1.3	3413	1	W52849	A. mediterranei ri
181	7	1.3	3729	1	W22603	Tyactone synthase
182	6	1.2	6	1	W19161	Isoelectric point
183	6	1.2	6	1	W18225	Purification tag o
184	6	1.2	6	1	W44011	Poly-His tag for c
185	6	1.2	6	1	W69961	Poly-His tag for c
186	6	1.2	6	1	W68297	Poly-His tag for c
187	6	1.2	6	1	W63024	Hexa-histidine tag
188	6	1.2	6	1	W84203	Peptide comprising
189	6	1.2	7	1	P80401	Affinity peptide c
190	6	1.2	7	1	Y03173	Human cardiac trop
191	6	1.2	8	1	R77412	Extension peptide
192	6	1.2	8	1	R95439	Polyhistidine tail
193	6	1.2	8	1	W20036	C-terminal metal b
194	6	1.2	8	1	W61479	A. fumigatus aller
195	6	1.2	9	1	R97378	Penta-histidine pe
196	6	1.2	9	1	W60080	Human sialins RH ep
197	6	1.2	9	1	W90199	B7.2-His tag fusi
198	6	1.2	10	1	R27517	Effector cell prot
199	6	1.2	10	1	R39733	First type QE-C-t
200	6	1.2	10	1	R73684	Metal-affinity dec
201	6	1.2	10	1	R74672	Polyhistidine moti
202	6	1.2	10	1	R77469	I-domain N-termi
203	6	1.2	10	1	W31496	Human DNase II pro
204	6	1.2	10	1	W56089	Human monocytic che
205	6	1.2	10	1	W76939	Fusion immunoglob
206	6	1.2	10	1	W82966	Human fibronectin
207	6	1.2	10	1	W82962	N-terminal hexahis
208	6	1.2	10	1	W94259	His-tag sequence.
209	6	1.2	10	1	Y04360	Salmonella flagell
210	6	1.2	11	1	R77410	Extension peptide
211	6	1.2	11	1	W53171	C-terminal peptide
212	6	1.2	11	1	W61480	A. fumigatus aller
213	6	1.2	11	1	W70581	Amnio acid sequen
214	6	1.2	11	1	W80426	Linker used in the
215	6	1.2	12	1	R07064	Transcript of plas
216	6	1.2	12	1	R45783	Farnesyltransferas
217	6	1.2	12	1	W61549	Endoprotease Xa
218	6	1.2	12	1	W82993	Human fibronectin
219	6	1.2	13	1	W61548	Endoprotease Xa
220	6	1.2	14	1	R60018	Can fit peptide vlt
221	6	1.2	14	1	W39836	Light chain CDRI O
222	6	1.2	14	1	W62982	Peptide inhibitor
223	6	1.2	15	1	R60513	Hexahistidine-cont
224	6	1.2	15	1	W72895	Mycobacterium tube
225	6	1.2	16	1	R69813	Cleavage site reco
226	6	1.2	16	1	W07945	gp120 peptide 17.
227	6	1.2	16	1	W30929	Birch pollen aller
228	6	1.2	17	1	P92081	Peptide coded for
229	6	1.2	17	1	R37708	Delta14 Ser17 hCNT
230	6	1.2	17	1	R97389	CC49 VH-spacer-PLA
231	6	1.2	17	1	W95000	N-terminal histidi
232	6	1.2	18	1	R60253	Vector-encoded NF-
233	6	1.2	18	1	R69814	Site recognised En
234	6	1.2	18	1	R78152	Synthetic leader t
235	6	1.2	18	1	R95870	CR2 receptor ligan
236	6	1.2	18	1	W71424	Peptide having an
237	6	1.2	19	1	W37707	C-terminal additio
238	6	1.2	19	1	W98067	Streptococcus pyog
239	6	1.2	20	1	R73014	N-terminal peptide
240	6	1.2	20	1	W00869	T-cell receptor al
241	6	1.2	20	1	W01445	Bioelastic polymer
242	6	1.2	20	1	W76938	Fusion immunoglob
243	6	1.2	21	1	R07063	Transcript of plas
244	6	1.2	21	1	P71703	Internal fragment
245	6	1.2	21	1	R60142	pGEX-6pPro proteas
246	6	1.2	21	1	W48442	Myc-his peptide ta
247	6	1.2	21	1	W54281	Myc-histidine pep
248	6	1.2	21	1	W54435	Human PsII2 expres
249	6	1.2	21	1	W75875	Myc-his tag peptid
250	6	1.2	21	1	W79683	Human CS198 protei
251	6	1.2	21	1	W86030	Myc-His peptide ta
252	6	1.2	21	1	W95563	Myc-his tag used i

253	6	1.2	21	1	W92257	Myc-his-tag sequen	326	6	1.2	94	1	R89747	Afr-1 interacting
254	6	1.2	21	1	W97661	Myc-his tag. Nucle	327	6	1.2	95	1	W77613	Deoxyribose-phosph
255	6	1.2	21	1	W95646	c-myc oncoprotein	328	6	1.2	97	1	W06408	Histidine tagged v
256	6	1.2	21	1	Y06793	Myc-His peptide ta	329	6	1.2	97	1	W06410	Histidine tagged v
257	6	1.2	21	1	Y07741	Human breast-speci	330	6	1.2	97	1	W06454	Recombinant human
258	6	1.2	22	1	W67267	Peptide #14 with c	331	6	1.2	99	1	W68490	Light chain of the
259	6	1.2	23	1	R79201	TGF-beta4 residues	332	6	1.2	100	1	W39892	Antigen 1 from clu
260	6	1.2	23	1	R94313	Histidine-6 insert	333	6	1.2	104	1	W89951	Haspnp polypeptide
261	6	1.2	23	1	W95472	Plasmid construct	334	6	1.2	105	1	W00920	Mycobacterium spec
262	6	1.2	25	1	R79206	TGF-beta4 residues	335	6	1.2	105	1	Y04772	Peptide inhibitor
263	6	1.2	25	1	W10250	Hepatitis B virus	336	6	1.2	107	1	W75892	Mycobacterium spec
264	6	1.2	25	1	W67272	Peptide #19 with c	337	6	1.2	107	1	Y04785	Peptide inhibitor
265	6	1.2	26	1	R64926	Four Corners Virus	338	6	1.2	108	1	R51502	Mycobacterium spec
266	6	1.2	26	1	W57233	Myelin basic protei	339	6	1.2	108	1	R51503	Pilin protein vari
267	6	1.2	28	1	W80575	N-terminal sequenc	340	6	1.2	108	1	Y04784	Mycobacterium spec
268	6	1.2	29	1	R65180	Portion of soluble	341	6	1.2	109	1	W39800	Variable domain of
269	6	1.2	30	1	W24708	Finger 2 domain of	342	6	1.2	109	1	Y00106	Enterococcus faeca
270	6	1.2	30	1	W65472	Hepatitis B surfac	343	6	1.2	112	1	R04027	X gene product of
271	6	1.2	31	1	W71427	Peptide having an	344	6	1.2	112	1	R34281	Human heavy chain
272	6	1.2	32	1	R97388	CC49 VH-spacer-TNF	345	6	1.2	112	1	W16469	Tyca peptide/carr
273	6	1.2	32	1	W71428	Peptide having an	346	6	1.2	113	1	W38475	S. pneumoniae. Nov
274	6	1.2	33	1	W71429	Peptide having an	347	6	1.2	116	1	W53994	Anti-CD4 antibody
275	6	1.2	33	1	W71445	Peptide having an	348	6	1.2	118	1	W08176	TGF-beta fusion pr
276	6	1.2	34	1	R37710	Delta14 Ser17 N-te	349	6	1.2	120	1	R42205	Valpha TCR gene w
277	6	1.2	34	1	W71430	Peptide having an	350	6	1.2	121	1	R42206	Vbeta TCR gene w
278	6	1.2	35	1	W71452	Peptide having an	351	6	1.2	122	1	W08181	TGF-beta fusion pr
279	6	1.2	35	1	W71453	Peptide having an	352	6	1.2	124	1	W08177	TGF-beta fusion pr
280	6	1.2	35	1	W71431	Peptide having an	353	6	1.2	124	1	W71693	R21-Hpro-28 fusion
281	6	1.2	36	1	W61376	GroEL N-terminal h	354	6	1.2	125	1	R80333	Protein polymeriz
282	6	1.2	36	1	W65068	E. coli GroEL N-te	355	6	1.2	126	1	R65292	HIV-1 fusion prote
283	6	1.2	36	1	W71432	Peptide having an	356	6	1.2	126	1	R90543	p164-5-CDK-BP clon
284	6	1.2	36	1	W71437	Peptide having an	357	6	1.2	128	1	W08188	TGF-beta fusion pr
285	6	1.2	36	1	W71438	Peptide having an	358	6	1.2	130	1	W02119	Vector expressing
286	6	1.2	37	1	W71433	Peptide having an	359	6	1.2	130	1	W08127	DE1-DE12-spacer-e
287	6	1.2	38	1	W71434	Peptide having an	360	6	1.2	130	1	W03981	TGF-beta fusion pr
288	6	1.2	40	1	W95473	PCANTAB vector enc	361	6	1.2	130	1	W08185	DE1-DE12-spacer-e
289	6	1.2	41	1	W34543	Artificial protein	362	6	1.2	130	1	W11285	TGF-beta fusion pr
290	6	1.2	43	1	R51891	scfv hexahistidine	363	6	1.2	130	1	W19623	Human src SH2 doma
291	6	1.2	43	1	W19897	Vector pUC119PHIS-	364	6	1.2	131	1	W08182	TGF-beta fusion pr
292	6	1.2	44	1	W95474	PCANTAB vector enc	365	6	1.2	132	1	R65290	Rat beta amyloid f
293	6	1.2	44	1	R75720	PELB signal sequen	366	6	1.2	132	1	R65291	Human beta amyloid
294	6	1.2	45	1	W04488	Staphylococcal ent	367	6	1.2	132	1	W08189	TGF-beta fusion pr
295	6	1.2	45	1	W73917	Staphylococcal ent	368	6	1.2	132	1	W75070	Human secreted pro
296	6	1.2	48	1	W06970	Modified hsp60 gen	369	6	1.2	133	1	W02121	DE1-DE12-spacer-e
297	6	1.2	48	1	Y00107	Enterococcus faeca	370	6	1.2	133	1	W03983	DE1-DE12-spacer-e
298	6	1.2	52	1	W80576	Human factor VII E	371	6	1.2	133	1	W20466	H. pylori transpor
299	6	1.2	54	1	W89178	Plasmid pOR4585 1n	372	6	1.2	133	1	W24668	TGF-beta fusion pr
300	6	1.2	55	1	R87580	Hepatitis G virus	373	6	1.2	133	1	W08178	DE1-DE12-spacer-e
301	6	1.2	55	1	W76092	HGV Q9 epitope pro	374	6	1.2	133	1	W11287	Human hcp SH2 doma
302	6	1.2	55	1	W80186	Peptide encoded by	375	6	1.2	133	1	W19625	Sequence of plant
303	6	1.2	55	1	W89481	Hepatitis G virus	376	6	1.2	134	1	P81139	Brassica seed acyl
304	6	1.2	57	1	W92793	US856134 Seq ID 2	377	6	1.2	134	1	R34979	B. campestris ACP.
305	6	1.2	57	1	W19696	Vector pUC119MCH-e	378	6	1.2	134	1	R76695	DE1-DE12-spacer-e
306	6	1.2	58	1	W99092	Oropouche NP prote	379	6	1.2	134	1	W02120	DE1-DE12-spacer-e
307	6	1.2	65	1	W94263	H6FXN12 fusion pr	380	6	1.2	134	1	W03982	ACP protein encode
308	6	1.2	67	1	R74528	Peptide encoded by	381	6	1.2	134	1	W15416	TGF-beta fusion pr
309	6	1.2	67	1	W80711	S. pneumoniae prot	382	6	1.2	134	1	W08187	TGF-beta fusion pr
310	6	1.2	68	1	Y12424	Human 5' EST seque	383	6	1.2	134	1	W08190	TGF-beta fusion pr
311	6	1.2	69	1	W94771	H6FXtr1pB fusion p	384	6	1.2	134	1	W11286	DE1-DE12-spacer-e
312	6	1.2	72	1	R88891	Measles virus F pr	385	6	1.2	134	1	W11289	DE1-DE12-spacer-e
313	6	1.2	72	1	W71692	Human defensin pre	386	6	1.2	134	1	W19624	Human lck SH2 doma
314	6	1.2	73	1	W94270	H6FXtr1pA fusion p	387	6	1.2	134	1	W19628	Human lck SH2 doma
315	6	1.2	73	1	W81027	Beta-gal IGF-1 fus	388	6	1.2	134	1	W30671	Brassica campestr
316	6	1.2	76	1	R13758	S. pneumoniae prot	389	6	1.2	134	1	W30527	Brassica campestr
317	6	1.2	80	1	Y11514	Mycobacterium spec	390	6	1.2	135	1	W08183	TGF-beta fusion pr
318	6	1.2	84	1	Y04775	Mouse dectin-2 His	391	6	1.2	136	1	W20850	H. pylori transpor
319	6	1.2	85	1	W69337	Multi-domain prote	392	6	1.2	137	1	W08186	TGF-beta fusion pr
320	6	1.2	87	1	R82953	Fibronectin fragme	393	6	1.2	137	1	P70101	Sequence encoded b
321	6	1.2	89	1	W19974	Human prostate pro	394	6	1.2	137	1	W08180	TGF-beta fusion pr
322	6	1.2	89	1	W23133	Fibronectin III-Q	395	6	1.2	137	1	W08184	Thermostable facto
323	6	1.2	89	1	W57195	Human fibronectin	396	6	1.2	139	1	W97107	TGF-beta fusion pr
324	6	1.2	89	1	W82985	CMV500-CREB2ZIP. N	397	6	1.2	139	1	W08179	
325	6	1.2	92	1	W00944		398	6	1.2	139	1	W31556	

399	6	1.2	143	1	R54047	Sequence of the VR
400	6	1.2	143	1	Y04827	Mycobacterium spec
401	6	1.2	144	1	R15178	hcg histidine subs
402	6	1.2	145	1	W11288	DET1-DET2-spacer-e
403	6	1.2	145	1	W19626	Human Stat6 SH2 do
404	6	1.2	145	1	W94266	H6XtRIPB-UB fusio
405	6	1.2	146	1	W42691	Recombinant rabbit
406	6	1.2	147	1	R27495	Fusion protein exp
407	6	1.2	152	1	W94265	H6FXTN3 fusion pro
408	6	1.2	157	1	W01774	1-153 (delta 112-1
409	6	1.2	158	1	W01774	(His)6/1-153 (delt
410	6	1.2	159	1	W31552	Collagen binding P
411	6	1.2	159	1	W69170	N-terminally tagged
412	6	1.2	162	1	W42692	Recombinant human
413	6	1.2	165	1	R42838	Prochymosin N-term
414	6	1.2	166	1	W59153	L. lactis rhbB C-t
415	6	1.2	166	1	W59162	L. lactis MG1316 r
416	6	1.2	171	1	R90765	CROC-1 c-fos promo
417	6	1.2	171	1	W46892	Amino acid sequenc
418	6	1.2	172	1	R37562	Human papilloma vi
419	6	1.2	172	1	R37563	Human papilloma vi
420	6	1.2	172	1	W65332	Trypanosoma cruzi
421	6	1.2	173	1	W60513	Antigenic determin
422	6	1.2	174	1	W06101	Proteolipid protel
423	6	1.2	174	1	W95083	S. aureus phosphor
424	6	1.2	174	1	W95084	Phosphoribosyl tra
425	6	1.2	177	1	W68589	Staphylococcus aur
426	6	1.2	178	1	R95259	ABC protease P II.
427	6	1.2	180	1	W69236	Mouse dectin-1 His
428	6	1.2	180	1	W94264	H6FXTN23 fusion pr
429	6	1.2	182	1	W72889	Mycobacterium tube
430	6	1.2	183	1	W06623	Sequence A encoded
431	6	1.2	185	1	W71501	Helicobacter polyp
432	6	1.2	192	1	R80345	Protein polymeric
433	6	1.2	195	1	R97229	Recombinant Brucel
434	6	1.2	196	1	R75335	Human oncostatin-M
435	6	1.2	197	1	W70495	Human disease rela
436	6	1.2	197	1	W94262	H6FXTN123 fusion P
437	6	1.2	198	1	R80346	Protein polymeric
438	6	1.2	198	1	W57672	Collagen-like poly
439	6	1.2	198	1	W49722	Protein polymer ad
440	6	1.2	201	1	W04573	HCV NS3 construct
441	6	1.2	201	1	W12966	HCV solubilised NS
442	6	1.2	201	1	W01644	HCV NS3 soluble pr
443	6	1.2	203	1	W06107	Foetal myelin basi
444	6	1.2	203	1	W06108	Foetal myelin basi
445	6	1.2	204	1	R07470	Transcript of plas
446	6	1.2	204	1	W14318	(His)6 modified CH
447	6	1.2	206	1	R33381	Cytokine SOSM. Hyb
448	6	1.2	206	1	W70222	Leishmania antigen
449	6	1.2	210	1	W04582	HCV NS3 protease c
450	6	1.2	210	1	W09241	HCV insoluble NS3
451	6	1.2	210	1	W01650	HCV NS3 protease c
452	6	1.2	211	1	R06501	GST-47 encoded pro
453	6	1.2	211	1	W31553	Collagen binding P
454	6	1.2	214	1	Y04472	Human Ras-like pro
455	6	1.2	217	1	R08347	Protein polymeric
456	6	1.2	218	1	W37339	Mouse HPRT. Multiv
457	6	1.2	218	1	W47019	Arabidopsis thalia
458	6	1.2	218	1	W47022	Arabidopsis thalia
459	6	1.2	218	1	W47018	Arabidopsis thalia
460	6	1.2	219	1	W23459	Recombinant MAM su
461	6	1.2	220	1	W47020	Arabidopsis thalia
462	6	1.2	223	1	R12899	Oncostatin M delet
463	6	1.2	223	1	R80348	protein polymeric
464	6	1.2	223	1	R80720	BLEF2 protein. Eps
465	6	1.2	223	1	W47351	Epstein-Barr virus
466	6	1.2	224	1	R12898	Oncostatin M delet
467	6	1.2	224	1	R55278	HTV2(gp33)-CD4 ex
468	6	1.2	226	1	W18054	Recombinant human
469	6	1.2	226	1	W72902	Mycobacterium tube
470	6	1.2	226	1	Y03168	Recombinant Human
471	6	1.2	226	1	Y03174	Recombinant human
472	6	1.2	227	1	P82309	Oncostatin M New D
473	6	1.2	227	1	R12871	Oncostatin M delet
474	6	1.2	227	1	R12896	Oncostatin M subtl
475	6	1.2	227	1	R12895	Oncostatin M subtl
476	6	1.2	227	1	R12897	Oncostatin M subtl
477	6	1.2	227	1	R12900	Oncostatin M subtl
478	6	1.2	227	1	R12901	Oncostatin M delet
479	6	1.2	227	1	W05530	Oncostatin M delet
480	6	1.2	228	1	W94261	Mouse TRADD intrac
481	6	1.2	230	1	R13593	CIH6FXTN123 fusio
482	6	1.2	233	1	R13303	Oncostatin M inser
483	6	1.2	233	1	R43011	Staphylococcal ent
484	6	1.2	233	1	W07613	Staphylococcal ent
485	6	1.2	233	1	W06738	Staphylococcal ent
486	6	1.2	233	1	W35373	Staphylococcal ent
487	6	1.2	233	1	W45516	Staphylococcus ent
488	6	1.2	233	1	W64570	NANC-2 heavy chal
489	6	1.2	234	1	W69607	Human UC PANCA mon
490	6	1.2	234	1	W69608	Human cysteine mod
491	6	1.2	235	1	R80349	Human cysteine mod
492	6	1.2	235	1	W69611	Protein polymeric
493	6	1.2	236	1	W69612	Human cysteine mod
494	6	1.2	242	1	W37700	Human cysteine mod
495	6	1.2	243	1	P70483	Protein sequence o
496	6	1.2	243	1	W37701	Protein sequence o
497	6	1.2	244	1	W07614	Ulcerative colitis
498	6	1.2	244	1	W64672	Human UC PANCA mon
499	6	1.2	244	1	W90309	hb7.1hs soluble f
500	6	1.2	245	1	W35375	Staphylococcus ent
501	6	1.2	246	1	W37703	Protein sequence o
502	6	1.2	247	1	R42382	H. somnus lipa. Ha
503	6	1.2	248	1	R42208	scvAlphavetabetaIb
504	6	1.2	248	1	W63830	Transformant CDW/1
505	6	1.2	250	1	R42307	scvAlphavetabeta with
506	6	1.2	250	1	W09239	HCV solubilised NS
507	6	1.2	250	1	W01648	HCV NS3 soluble pr
508	6	1.2	250	1	W14944	Chimeric human/por
509	6	1.2	251	1	W49735	Protein polymer ad
510	6	1.2	252	1	P92082	Consensus sequence
511	6	1.2	252	1	R33380	Cytokine hOSM. Hyb
512	6	1.2	252	1	R88202	Human oncostatin M
513	6	1.2	253	1	P94685	Amino acid sequenc
514	6	1.2	253	1	W80667	S. pneumoniae lacra
515	6	1.2	253	1	W90225	Anti-B7.1 monosp
516	6	1.2	254	1	W41996	Flea cysteine prot
517	6	1.2	254	1	W41997	HCV NS3-NS4A prote
518	6	1.2	255	1	W09237	HCV solubilised NS
519	6	1.2	255	1	W01646	HCV NS3 soluble pr
520	6	1.2	255	1	W37702	Protein sequence o
521	6	1.2	255	1	W47147	Chloramphenicol ac
522	6	1.2	255	1	W54875	Staphylococcus aur
523	6	1.2	256	1	W90221	Anti-B7.1 monosp
524	6	1.2	258	1	R77617	Anti-c5 Mab N19/8
525	6	1.2	260	1	R82955	Three domain prote
526	6	1.2	261	1	R97726	B10 single chain T
527	6	1.2	262	1	W49734	Protein polymer ad
528	6	1.2	262	1	P93565	Sequence of an epi
529	6	1.2	263	1	P90507	Sequence of an epi
530	6	1.2	263	1	R82957	Three domain prote
531	6	1.2	263	1	R82957	D10 single chain T
532	6	1.2	263	1	R95968	Nucleic acid recog
533	6	1.2	263	1	R95968	Nucleic acid recog
534	6	1.2	263	1	R95969	Nucleic acid recog
535	6	1.2	263	1	W32447	Mycobacterium tube
536	6	1.2	263	1	W32379	Mycobacterium tube
537	6	1.2	263	1	W64317	Mycobacterium tube
538	6	1.2	263	1	W81680	Anti-B7.2 monosp
539	6	1.2	263	1	W90226	Anti-B7.2 monosp
540	6	1.2	264	1	R81428	Hepatitis GB virus
541	6	1.2	266	1	P91236	(Hexahis)(mdhFR)(E
542	6	1.2	266	1	R97561	Human papilloma vi
543	6	1.2	268	1	W61207	Streptococcus pneu
544	6	1.2	268	1	W90222	Anti-B7.2 monosp

545	6	1.2	269	1	R40801	618	6	1.2	343	1	R82455	Hepatitis GB virus
546	6	1.2	270	1	W04571	619	6	1.2	343	1	W60053	Human protein tyro
547	6	1.2	270	1	W09236	620	6	1.2	344	1	W08097	Vmpt soluble varia
548	6	1.2	270	1	W01645	621	6	1.2	345	1	W18879	Leucine aminopepti
549	6	1.2	272	1	W26698	622	6	1.2	345	1	W15555	Fibronectin-bindin
550	6	1.2	272	1	W87637	623	6	1.2	346	1	R39493	Human apoAIV mutel
551	6	1.2	274	1	W39899	624	6	1.2	348	1	P70593	Human apoAIV mutel
552	6	1.2	275	1	R84563	625	6	1.2	349	1	R43911	Fructose-bisphosph
553	6	1.2	276	1	W90227	626	6	1.2	349	1	W06182	Sequence encoded b
554	6	1.2	277	1	P70298	627	6	1.2	349	1	W06173	Modified HCPB (D25
555	6	1.2	277	1	W33275	628	6	1.2	349	1	W06181	mature HCPB-(H195
556	6	1.2	281	1	W49736	629	6	1.2	349	1	W13750	Modified HCPB (D25
557	6	1.2	281	1	W82744	630	6	1.2	349	1	W13751	Carboxypeptidase B
558	6	1.2	281	1	W90223	631	6	1.2	349	1	W13752	Carboxypeptidase B
559	6	1.2	282	1	Y13365	632	6	1.2	349	1	W13753	Carboxypeptidase B
560	6	1.2	284	1	W64366	633	6	1.2	349	1	W13757	Carboxypeptidase B
561	6	1.2	284	1	W81733	634	6	1.2	349	1	W13758	Carboxypeptidase B
562	6	1.2	286	1	R49836	635	6	1.2	349	1	W13759	Carboxypeptidase B
563	6	1.2	286	1	W98520	636	6	1.2	349	1	W13760	Carboxypeptidase B
564	6	1.2	287	1	W49737	637	6	1.2	349	1	W13761	Carboxypeptidase B
565	6	1.2	288	1	W82316	638	6	1.2	349	1	W13765	Carboxypeptidase B
566	6	1.2	288	1	W82317	639	6	1.2	349	1	W13763	Carboxypeptidase B
567	6	1.2	288	1	W82482	640	6	1.2	349	1	W13765	Carboxypeptidase B
568	6	1.2	288	1	W82743	641	6	1.2	349	1	W13765	Carboxypeptidase B
569	6	1.2	291	1	W49738	642	6	1.2	349	1	W13749	Carboxypeptidase B
570	6	1.2	297	1	R56543	643	6	1.2	351	1	W56000	Fibroblast growth
571	6	1.2	297	1	W15008	644	6	1.2	351	1	R81355	Thrombospondin pro
572	6	1.2	301	1	W1507	645	6	1.2	353	1	R81355	Thrombospondin pro
573	6	1.2	301	1	W73217	646	6	1.2	353	1	R81356	Thrombospondin pro
574	6	1.2	303	1	W08956	647	6	1.2	355	1	Y06997	Galectin-9 protein
575	6	1.2	304	1	R73599	648	6	1.2	356	1	W22676	Borrelia variable
576	6	1.2	308	1	W13322	649	6	1.2	357	1	R66810	Extracellular doma
577	6	1.2	308	1	W34322	650	6	1.2	359	1	W61046	His-tagged polylin
578	6	1.2	309	1	W83322	651	6	1.2	360	1	R75006	Human GSBP2. New C
579	6	1.2	310	1	W83324	652	6	1.2	360	1	W37908	Tomato S-adenosyl-
580	6	1.2	311	1	W56504	653	6	1.2	360	1	W49066	Tbv25-Pfs28C fusio
581	6	1.2	312	1	W05528	654	6	1.2	360	1	W49825	Cytokine suppress
582	6	1.2	312	1	W83323	655	6	1.2	363	1	W37910	Human monocytic CSA
583	6	1.2	315	1	R96268	656	6	1.2	365	1	P94144	Tbv25-28 fusion pr
584	6	1.2	316	1	Y01820	657	6	1.2	365	1	R05080	Prothymosin. DNA I
585	6	1.2	317	1	W61195	658	6	1.2	365	1	P30603	Sequence of calf p
586	6	1.2	319	1	W35855	659	6	1.2	365	1	W99874	Sequence encoded b
587	6	1.2	320	1	W24210	660	6	1.2	368	1	W01808	Case-A2 VI/V2 fusi
588	6	1.2	320	1	W38217	661	6	1.2	369	1	W98612	Human papillomavir
589	6	1.2	320	1	W59937	662	6	1.2	369	1	Y07080	H. pylori GHPD 202
590	6	1.2	320	1	W70965	663	6	1.2	372	1	W82481	Renal cancer assoc
591	6	1.2	323	1	W06407	664	6	1.2	373	1	R39486	Mammalian Zap-70 p
592	6	1.2	326	1	W06409	665	6	1.2	373	1	W06103	Human apoAIV mutel
593	6	1.2	329	1	W44778	666	6	1.2	373	1	Y07203	MP4 chimera (MBP21
594	6	1.2	330	1	W94267	667	6	1.2	374	1	W42434	Mycobacterium spec
595	6	1.2	331	1	P70328	668	6	1.2	375	1	P40078	Murine promiscuous
596	6	1.2	331	1	W94268	669	6	1.2	375	1	W06104	Sequence encoded b
597	6	1.2	332	1	W14919	670	6	1.2	375	1	W01809	PM4 chimera (delta
598	6	1.2	332	1	W32418	671	6	1.2	376	1	R21416	Human papillomavir
599	6	1.2	332	1	W32350	672	6	1.2	377	1	R21417	Carbolic anhydraz
600	6	1.2	332	1	W64322	673	6	1.2	377	1	W05589	Chlamydomonas carb
601	6	1.2	332	1	W76189	674	6	1.2	377	1	W05589	Aminopeptidase. En
602	6	1.2	332	1	W59951	675	6	1.2	377	1	W86195	Human fc receptor
603	6	1.2	333	1	W81683	676	6	1.2	378	1	R50076	NANBH virus antige
604	6	1.2	333	1	R31019	677	6	1.2	378	1	W09219	SELPK polymer. T1
605	6	1.2	334	1	W98714	678	6	1.2	378	1	W53545	Amino acid sequenc
606	6	1.2	337	1	P81191	679	6	1.2	378	1	W71072	Multiple sclerosis
607	6	1.2	337	1	R39485	680	6	1.2	379	1	W98558	Protein encoded by
608	6	1.2	338	1	R80350	681	6	1.2	379	1	P90508	Sequence of an epi
609	6	1.2	339	1	W37909	682	6	1.2	380	1	P30013	Sequence encoded b
610	6	1.2	341	1	W37907	683	6	1.2	380	1	P90510	Sequence of an epi
611	6	1.2	342	1	R39487	684	6	1.2	380	1	P90511	Sequence of an epi
612	6	1.2	342	1	R39489	685	6	1.2	380	1	P90512	Sequence of new po
613	6	1.2	342	1	R39491	686	6	1.2	380	1	P20038	Pre-prorenin-A pr
614	6	1.2	342	1	R39496	687	6	1.2	380	1	R37714	hCNTF fusion proti
615	6	1.2	342	1	R39498	688	6	1.2	381	1	W19111	Human tumour succe
616	6	1.2	342	1	R95054	689	6	1.2	381	1	R20730	Prochymosin (prote
617	6	1.2	342	1	W55958	690	6	1.2	381	1	P40218	Sequence of rennin

691 6 1.2 381 1 P40559
 692 6 1.2 381 1 Sequence encoded b
 693 6 1.2 381 1 Sequence encoded b
 694 6 1.2 381 1 Mouse tumour susce
 695 6 1.2 381 1 Mouse tsg101 prote
 696 6 1.2 383 1 Bruce1a abortus 3
 697 6 1.2 386 1 Cyclin E (MRHHNH)
 698 6 1.2 386 1 Globodera rostocki
 699 6 1.2 386 1 Human TNF receptor
 700 6 1.2 386 1 Human TNF receptor
 701 6 1.2 386 1 Human TRAIL recept
 702 6 1.2 386 1 Human TANGO-74 pro
 703 6 1.2 386 1 RTD, inhibitor of
 704 6 1.2 390 1 New sarcosine oxid
 705 6 1.2 390 1 Bacillus sarcosine
 706 6 1.2 390 1 Human NHR5 protein
 707 6 1.2 391 1 Soybean seed stea
 708 6 1.2 391 1 Stearoyl-ACP-desat
 709 6 1.2 391 1 Single Chain T cel
 710 6 1.2 391 1 Mycobacterium tube
 711 6 1.2 391 1 Mycobacterium tube
 712 6 1.2 391 1 Amino acid sequenc
 713 6 1.2 391 1 Mycobacterium tube
 714 6 1.2 391 1 M. tuberculosis Im
 715 6 1.2 391 1 Mycobacterium spec
 716 6 1.2 392 1 Enterococcus faeca
 717 6 1.2 394 1 Mycobacterium spec
 718 6 1.2 395 1 Transcription fact
 719 6 1.2 397 1 Human O-fucosyltra
 720 6 1.2 397 1 Protein encoded by
 721 6 1.2 400 1 Sequence of prepro
 722 6 1.2 400 1 H. pylori ORF 06p
 723 6 1.2 406 1 Thyroid hormone re
 724 6 1.2 406 1 Human 26S proteaso
 725 6 1.2 406 1 Thyroid hormone re
 726 6 1.2 406 1 Human JLI protein.
 727 6 1.2 407 1 Flock house virus
 728 6 1.2 407 1 FHV capsid protein
 729 6 1.2 409 1 FHV capsid protein
 730 6 1.2 409 1 Tobacco laccase cl
 731 6 1.2 410 1 FHV capsid protein
 732 6 1.2 415 1 FHV capsid protein
 733 6 1.2 418 1 Soybean SUG1 polyp
 734 6 1.2 421 1 Rat intrinsic fact
 735 6 1.2 421 1 Rat intrinsic fact
 736 6 1.2 421 1 IL-2-DETA-DGALA mu
 737 6 1.2 423 1 An Escherichia col
 738 6 1.2 425 1 Fragment of the hu
 739 6 1.2 432 1 Carboxypeptidase C
 740 6 1.2 436 1 Threonine dehydrat
 741 6 1.2 436 1 Threonine dehydrat
 742 6 1.2 436 1 Threonine dehydrat
 743 6 1.2 436 1 Threonine dehydrat
 744 6 1.2 436 1 Threonine dehydrat
 745 6 1.2 436 1 Threonine dehydrat
 746 6 1.2 436 1 Threonine dehydrat
 747 6 1.2 441 1 Nucleotide sequenc
 748 6 1.2 443 1 Mouse H74 protein.
 749 6 1.2 445 1 NADH:H2O oxidase a
 750 6 1.2 445 1 Clostridium botuli
 751 6 1.2 446 1 Staphylococcus aur
 752 6 1.2 446 1 Bacterial and mamm
 753 6 1.2 450 1 Sequence encoded b
 754 6 1.2 450 1 Tyrosine kinase as
 755 6 1.2 453 1 Aspergillus aculea
 756 6 1.2 457 1 Acromonium cellulo
 757 6 1.2 458 1 BamHI/SalI insert
 758 6 1.2 459 1 Vector SSCI-derivate
 759 6 1.2 459 1 SSCI single chain
 760 6 1.2 463 1 TGF alpha-ETA fusi
 761 6 1.2 465 1 Human papillomavir
 762 6 1.2 469 1 B. flavum fumatease
 763 6 1.2 472 1 Globodera rostocki
 6 1.2 472 1 Kidney injury asso

764 6 1.2 474 1 W36132
 765 6 1.2 477 1 R43999
 766 6 1.2 479 1 R49739
 767 6 1.2 482 1 Y06915
 768 6 1.2 481 1 Y06915
 769 6 1.2 481 1 W13373
 770 6 1.2 482 1 W06105
 771 6 1.2 483 1 R13992
 772 6 1.2 484 1 R90765
 773 6 1.2 494 1 W42996
 774 6 1.2 495 1 W37062
 775 6 1.2 495 1 W37062
 776 6 1.2 496 1 R95058
 777 6 1.2 498 1 R28805
 778 6 1.2 511 1 W78475
 779 6 1.2 512 1 W31554
 780 6 1.2 514 1 W31876
 781 6 1.2 514 1 W83359
 782 6 1.2 515 1 W88426
 783 6 1.2 515 1 P91235
 784 6 1.2 519 1 W00181
 785 6 1.2 521 1 R69997
 786 6 1.2 524 1 R56525
 787 6 1.2 524 1 R76621
 788 6 1.2 525 1 R99466
 789 6 1.2 527 1 R82954
 790 6 1.2 530 1 R95053
 791 6 1.2 530 1 W28501
 792 6 1.2 530 1 W47126
 793 6 1.2 530 1 W78476
 794 6 1.2 532 1 R42457
 795 6 1.2 532 1 R83016
 796 6 1.2 533 1 W19763
 797 6 1.2 537 1 Y03839
 798 6 1.2 539 1 W48358
 799 6 1.2 544 1 W46314
 800 6 1.2 546 1 W38234
 801 6 1.2 547 1 W22707
 802 6 1.2 550 1 R42085
 803 6 1.2 550 1 R42254
 804 6 1.2 550 1 R42255
 805 6 1.2 551 1 W38238
 806 6 1.2 552 1 W48873
 807 6 1.2 552 1 W96299
 808 6 1.2 553 1 W22706
 809 6 1.2 553 1 W1508
 810 6 1.2 554 1 W73223
 811 6 1.2 554 1 R70827
 812 6 1.2 554 1 W56099
 813 6 1.2 556 1 W98849
 814 6 1.2 556 1 W90218
 815 6 1.2 559 1 W78753
 816 6 1.2 561 1 R48058
 817 6 1.2 564 1 W48788
 818 6 1.2 566 1 P82595
 819 6 1.2 566 1 R08263
 820 6 1.2 572 1 Y04954
 821 6 1.2 573 1 W48874
 822 6 1.2 575 1 R11867
 823 6 1.2 580 1 W90217
 824 6 1.2 585 1 W31871
 825 6 1.2 591 1 W21963
 826 6 1.2 592 1 W38233
 827 6 1.2 592 1 W78754
 828 6 1.2 592 1 W94269
 829 6 1.2 593 1 R86292
 830 6 1.2 596 1 W48785
 831 6 1.2 597 1 W21964
 832 6 1.2 598 1 W46313
 833 6 1.2 601 1 W03562
 834 6 1.2 601 1 W44291
 835 6 1.2 604 1 P60359
 836 6 1.2 606 1 W85019
 6 1.2 607 1 W38237

N-terminal modifie
 Human OTC. Gene mu
 Protein polymer ad
 Amino acid sequenc
 Recombinant Sp 100
 MmOG4 chimera (MS
 P. falciparum spor
 Human K+ channel 2
 Putative mature po
 HIV-1 breakthrough
 HIV-1 breakthrough
 GAL4-DT-IL-2 multi
 Soluble HIV gp120.
 Autographa callifor
 Collagen binding p
 GFP variants S65C
 Streptococcus pneu
 Chlamydia pneumonia
 (ENV-80)(GAG-VII)(
 HIV-1 env protein.
 HIV-1 env protein.
 Protein kinase (NU
 Saccharomyces cere
 Biosynthetic enzym
 Six domain protein
 scFv(FRP5)-DETA-DG
 Birch pollen co-fa
 Uridine diphospho
 Baculovirus ISF pr
 Enzyme involved in
 Recombinant papill
 p53-GM-CSF immuno
 Chimeric transcrip
 Fractionated X irr
 AAV4 vps coat prot
 VEGF/CPG2 fusion p
 N-terminal truncat
 NF-KappaB p65 prot
 Human p65 protein
 Human p65 protein
 VEGF/CPG2 fusion p
 Hypophyoma sp. stra
 Vgag streptogramin
 Arginyl-tRNA synth
 Single chain anti-
 H22-anti-CRA antib
 Anti-catalect immu
 A. oryzae ATCC2038
 H. pylori GHP 160
 Bispecific tetrava
 Human EYAL-A. Nuci
 Sequence of protea
 Thyroid peroxidase
 Amylase from strep
 Mycobacterium spec
 Hypophyoma sp. stra
 Delta-endotoxin P
 Bispecific tetrava
 Human fitzied-5 p
 Recombinant furin
 VEGF/CPG2 fusion p
 Human EYAL-B. Nuci
 H6FSCPV(CRA6)trip
 Pseudomonas mesoac
 Thyroid peroxidase
 Recombinant furin
 AAV4 vps coat prot
 Human papillomavir
 Human papillomavir
 Glucoamylase struc
 p38-green fluoresc
 VEGF/CPG2 fusion p

837	6	1.2	607	1	W85010	p38-green fluoresc	910	6	1.2	797	1	W85042	p38-green fluoresc
838	6	1.2	609	1	W82726	Adenovirus PAC75S1	911	6	1.2	797	1	W85043	NFKappaB p65 subun
839	6	1.2	609	1	W82727	Adenovirus p15/gN	912	6	1.2	798	1	W06592	Amino acid sequenc
840	6	1.2	610	1	R10923	Mouse protein kin	913	6	1.2	798	1	W38188	Aradidopsis SCAREC
841	6	1.2	612	1	W09321	Human mast cell-de	914	6	1.2	811	1	R41867	Mouse OSF-2. Mamma
842	6	1.2	615	1	R95056	GAL4-DT-scfv(FRP5)	915	6	1.2	816	1	R39647	Human fibroblast g
843	6	1.2	616	1	W41784	GST-cyclin E-PEST+	916	6	1.2	819	1	R10648	Human basic fibr
844	6	1.2	617	1	R95057	GAL4-DT-scfv(FRP5)	917	6	1.2	820	1	R47233	Human fibroblast g
845	6	1.2	620	1	W14994	Human c-Fos induce	918	6	1.2	820	1	R71802	N-acetylheparosan
846	6	1.2	624	1	W48789	Thyroid peroxidase	919	6	1.2	820	1	W63844	Mouse DGF recepto
847	6	1.2	624	1	W98336	H. pylori GHP0 139	920	6	1.2	822	1	R13549	Basic FGF receptor
848	6	1.2	628	1	R64579	Pig Syk. Chimeric	921	6	1.2	822	1	R20750	flg receptor prote
849	6	1.2	630	1	R62868	Tyrosine-kinase Sy	922	6	1.2	822	1	R26337	N-sam. Gene prod o
850	6	1.2	631	1	W23036	Human cytomagalovi	923	6	1.2	829	1	R80334	Protein polymERIC
851	6	1.2	634	1	W98888	Antigen from clust	924	6	1.2	829	1	W57673	Collagen-like poly
852	6	1.2	637	1	P92108	Human growth hormo	925	6	1.2	829	1	W49723	Protein polymer ad
853	6	1.2	638	1	P81326	Human growth hormo	926	6	1.2	832	1	W04302	Antibody/beta gluc
854	6	1.2	638	1	W33594	Human growth hormo	927	6	1.2	832	1	R41869	Human OSF-2. Mamma
855	6	1.2	651	1	W05135	scFv(225)-ETA fusi	928	6	1.2	846	1	R93804	Class A starch bra
856	6	1.2	651	1	W05136	scFv(FRP5)-ETA fus	929	6	1.2	847	1	W06400	Class A starch bra
857	6	1.2	653	1	W57437	NAML 0 fusion cons	930	6	1.2	848	1	Y07733	Human thyroid peto
858	6	1.2	662	1	W18572	Human cyclin D1-hu	931	6	1.2	852	1	W48782	Thyroid peroxidase
859	6	1.2	662	1	W74573	Human cyclin D1/cy	932	6	1.2	853	1	W43066	HIV-1 gp120 protel
860	6	1.2	666	1	W53063	Streptococcus pneu	933	6	1.2	856	1	R41025	Selectively deqlyc
861	6	1.2	668	1	R10686	Feline calcivir	934	6	1.2	856	1	R41025	Selectively deqlyc
862	6	1.2	671	1	W53458	Fly latheo protein	935	6	1.2	856	1	R41025	Selectively deqlyc
863	6	1.2	677	1	W55984	Swinepox virus str	936	6	1.2	856	1	R41027	Selectively deqlyc
864	6	1.2	677	1	W70900	Human beta galacto	937	6	1.2	856	1	R41028	Selectively deqlyc
865	6	1.2	681	1	W48786	Thyroid peroxidase	938	6	1.2	856	1	R41029	Selectively deqlyc
866	6	1.2	682	1	R80330	S. pneumoniae pent	939	6	1.2	856	1	R41030	Selectively deqlyc
867	6	1.2	682	1	W04359	Collagen-like poly	940	6	1.2	856	1	R41031	Selectively deqlyc
868	6	1.2	682	1	W57670	Protein polymer ad	941	6	1.2	856	1	R41032	Selectively deqlyc
869	6	1.2	682	1	W49717	Thyroid peroxidase	942	6	1.2	856	1	W82480	Mammalian Zap-70 p
870	6	1.2	689	1	W48787	scFv(FRP5)/TGF alp	943	6	1.2	859	1	W70796	Human gp130-Fc-His
871	6	1.2	689	1	W05138	scFv(FRP5)/TGF alp	944	6	1.2	863	1	R43873	HTLV-III ENV-LOR g
872	6	1.2	705	1	W18573	Human cyclin D1-hu	945	6	1.2	865	1	W85027	Zap70-green fluore
873	6	1.2	705	1	W74574	Human cyclin D1/cy	946	6	1.2	868	1	R07454	Second open readin
874	6	1.2	706	1	R75647	Thermophilic bacte	947	6	1.2	870	1	W17080	Clarkia breweri S-
875	6	1.2	712	1	R95648	Thermostable DNA-1	948	6	1.2	870	1	W73485	LinAlool synthase
876	6	1.2	719	1	W21961	Recombinant furin	949	6	1.2	871	1	W85026	Green fluorescent
877	6	1.2	711	1	R10649	Chicken basic fibr	950	6	1.2	871	1	W27658	Streptococcus pneu
878	6	1.2	731	1	R21686	Human bFGF recepto	951	6	1.2	877	1	R80335	Protein polymERIC
879	6	1.2	734	1	W46308	AAV4 VPI capsid pr	952	6	1.2	877	1	W49724	Protein polymer ad
880	6	1.2	737	1	R05270	Amino acid sequenc	953	6	1.2	878	1	R55060	Sequence of human
881	6	1.2	737	1	R56494	TATRA-binding prote	954	6	1.2	878	1	R85487	Human E-cadherin p
882	6	1.2	737	1	W06084	Human TATRA-binding	955	6	1.2	878	1	W19113	Potato starch bran
883	6	1.2	737	1	W25019	Mouse Fas-binding	956	6	1.2	880	1	W72093	HSV-2 strain SB5 C
884	6	1.2	739	1	W61532	Thyroid peroxidase	957	6	1.2	881	1	W48791	HSV-2 strain SB5 C
885	6	1.2	740	1	W48790	Helicobacter-speci	958	6	1.2	881	1	W72002	HSV-2 strain SB5 C
886	6	1.2	741	1	W06712	N. meningitidis st	959	6	1.2	881	1	W72174	Class A starch bra
887	6	1.2	753	1	W93495	Lucifilla cuprina ec	960	6	1.2	882	1	W06539	Protein polymer ad
888	6	1.2	757	1	R71297	Protein polymERIC	961	6	1.2	884	1	R80341	SELPRK polymer. ti
889	6	1.2	762	1	R80327	Amino acid sequenc	962	6	1.2	884	1	W09213	Expected amino aci
890	6	1.2	762	1	R80324	Collagen-like poly	963	6	1.2	884	1	W53541	SELPRK polymer. Ne
891	6	1.2	762	1	W53535	Protein polymer ad	964	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
892	6	1.2	762	1	W57666	Protein polymer ad	965	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
893	6	1.2	762	1	W57666	Protein polymer ad	966	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
894	6	1.2	762	1	W49713	Down-regulated in	967	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
895	6	1.2	762	1	W49715	Human placenta acu	968	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
896	6	1.2	770	1	R60568	Mouse liver acute	969	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
897	6	1.2	770	1	R72082	Mouse STAT3 allel1	970	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
898	6	1.2	770	1	R82993	Human STAT3 allel1	971	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
899	6	1.2	770	1	R82993	Human STAT3 allel1	972	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
900	6	1.2	770	1	W03176	Human STAT3 allel1	973	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
901	6	1.2	770	1	Y03768	Human STAT3 allel1	974	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
902	6	1.2	779	1	R74302	TCI protein. New t	975	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
903	6	1.2	779	1	R41688	Enzyme involved in	976	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
904	6	1.2	780	1	R42454	Biosynthetic enzym	977	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
905	6	1.2	780	1	R99463	S. putrefaciens EP	978	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
906	6	1.2	780	1	W37051	C-proteinase sequ	979	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
907	6	1.2	788	1	W75919	Bacillus thuringie	980	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
908	6	1.2	789	1	W46868	Mouse gamma TI ada	981	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET
909	6	1.2	791	1	W38331	Mouse gamma TI ada	982	6	1.2	884	1	W49728	scFv2(FRP5/225)-ET

983 1.2 942 1 W56699 Modified protein K
 984 1.2 948 1 W48781 Thyroid peroxidase
 985 1.2 956 1 R15784 B.thuringiensis to
 986 1.2 966 1 W09223 SERP-Csz polymer
 987 1.2 970 1 R85089 EPR-like receptor
 988 1.2 974 1 W64221 Human secreted pro
 989 1.2 976 1 W02289 Mouse neuron resti
 990 1.2 980 1 R32359 Human KA-2 recepto
 991 1.2 985 1 R42995 Glycosyltransferas
 992 1.2 985 1 R42214 Aspergillus niger
 993 1.2 986 1 W13670 C-proteinase encod
 994 1.2 990 1 R39343 EpiB protein. Nove
 995 1.2 994 1 R87018 Receptor. tyrosine
 996 1.2 994 1 W26366 Mouse Nuk tyrosine
 997 1.2 995 1 R75712 Eph-related PTK Ce
 998 1.2 995 1 R84658 Stripan-P pol prot
 999 1.2 1002 1 W09218 SERP-C polymer. T1
 1000 1.2 1002 1 W53544 Amino acid sequenc

ALIGNMENTS

RESULT 1
 R7961 ID R79961 standard; Protein; 329 AA.
 AC R79961; 18-APR-1996 (first entry)
 DE Partial lcrv (V antigen) of Y. pestis.
 KW lcrv; V antigen; virulence; plague; vaccine; epitope.
 OS Yersinia pestis.
 PN W09524475-A1.
 PD 14-SEP-1995.
 PR 06-MAR-1995; G00481.
 PR 08-MAR-1994; GB-004577.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Leary SEC, Titball RW, Williamson ED, Leary SE;
 DR WPI: 95-328268/42.
 DR N-PSDB: T04222.
 PR Recombinant DNA expressing Yersinia pestis V antigen - useful in
 PT oral or parenteral vaccines for protection against plague
 PS Claim 6: page 11-13; 25pp, English.
 CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrv) encoding
 CC all or a protective epitopic part of the mature V protein of Yersinia
 CC pestis. The protein was expressed as a fusion protein with maltose
 CC binding protein or glutathione-S-transferase in 3 different plasmid
 CC vectors. Y. pestis is the highly virulent causative organism of plague
 CC in a wide range of animals, including man. The V antigen (lcrv) is an
 CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
 CC The V antigen is postulated to act as a virulence antigen, and
 CC transformed microorganisms contg. recombinant DNA encoding a V antigen
 CC protein/peptide are useful in vaccines to protect against plague.
 SQ Sequence 329 AA;

Query Match 46.6%; Score 243; DB 1; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.2e-225;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHYDNOLONGIKRYKEFLLESSPNTOWELRAFMVHFSITADRIDDILKVIYDSM 338
 DB 87 ILKGGHYDNOLONGIKRYKEFLLESSPNTOWELRAFMVHFSITADRIDDILKVIYDSM 146
 QY 339 NHGDAKSLREELAEIETAEIKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 398
 DB 147 NHGDAKSLREELAEIETAEIKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 206
 QY 399 EEIFKASAEYKILEKMPOTTIOVDGSEKKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 458
 DB 207 EEIFKASAEYKILEKMPOTTIOVDGSEKKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 266
 QY 459 NELSHFATTCGSKSRPLNDLVSOQTTOISDITSRFNSAIEALNRFIOKYDSVMORLLDDT 518

DB 267 NELSHFATTCGSKSRPLNDLVSOQTTOISDITSRFNSAIEALNRFIOKYDSVMORLLDDT 326
 QY 519 SGK 521
 DB 327 SGK 329

RESULT 2
 W01040 ID W01040 standard; Protein; 329 AA.
 AC W01040; 28-DEC-1996 (first entry)
 DE Y. pestis V antigen.
 KW Plague; vaccine; genetic immunisation; V antigen; lcrv;
 OS Yersinia pestis strain GB.
 FH Key location/Qualifiers
 FT peptide 1..4
 FT /note= "vector-encoded peptide"
 PN W09628551-A1.
 PD 19-SEP-1996.
 PR 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR N-PSDB: T38242.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Disclosure: page 25-28; 98pp; English.
 CC Yersinia pestis V antigen (W01040) is capable of evoking protective
 CC immune responses in animals. A combined vaccine of V and F1
 CC antigens (see also W01042) can at least match the protection
 CC afforded by live attenuated EV76 vaccine without any of the hazards
 CC that have kept the EV vaccine from general use. The V antigen is
 CC produced using the lcrv gene (see also T38242) obt. from Y.
 CC pestis by PCR amplification. It can also be prep. as a fusion with
 CC F1 antigen (see also W01044-45) and expressed by attenuated Arco or
 CC C Salmonella typhi as a live vaccine for long-term protection
 CC against plague. Expression by gut-colonising bacterial transformants
 CC produces a protective response against Y. pestis.
 SQ Sequence 329 AA;

Query Match 46.6%; Score 243; DB 1; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.2e-225;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHYDNOLONGIKRYKEFLLESSPNTOWELRAFMVHFSITADRIDDILKVIYDSM 338
 DB 87 ILKGGHYDNOLONGIKRYKEFLLESSPNTOWELRAFMVHFSITADRIDDILKVIYDSM 146
 QY 339 NHGDAKSLREELAEIETAEIKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 398
 DB 147 NHGDAKSLREELAEIETAEIKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 206
 QY 399 EEIFKASAEYKILEKMPOTTIOVDGSEKKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 458
 DB 207 EEIFKASAEYKILEKMPOTTIOVDGSEKKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 266
 QY 459 NELSHFATTCGSKSRPLNDLVSOQTTOISDITSRFNSAIEALNRFIOKYDSVMORLLDDT 518
 DB 267 NELSHFATTCGSKSRPLNDLVSOQTTOISDITSRFNSAIEALNRFIOKYDSVMORLLDDT 326
 QY 519 SGK 521
 DB 327 SGK 329

RESULT 3
 W01041

ID W01041 standard; Protein: 329 AA.
AC W01041;
DE 28-DEC-1996 (first entry)
DE Y. pestis V antigen.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW Y. pestis V antigen.
OS Yersinia pestis strain GB.
FT Key
FT Location/Qualifiers
FT peptide 1..4
FT /note="vector-encoded peptide"
PN W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI; 96-433824/43.
DR N-PSDB; T38243.
PT Yersinia pestis V antigen and F1 antigen or their protective
PS epitopic parts - useful in vaccine for protection against plague
PS Disclosure; Page 32-35; 98pp; English.
CC Yersinia pestis V antigen (W01041) is capable of evoking protective
CC immune responses in animals. A combined vaccine of V and F1
CC antigens (see also W01042) can at least match the protection
CC afforded by live attenuated EV76 vaccine without any of the hazards
CC that have kept the lcrV gene (see also T38243) out of the V antigen is
CC produced using the lcrV gene (see also T38243) obt'd. from Y. pestis
CC by PCR amplification. It can also be prep'd. as a fusion with F1
CC antigen (see also W01044-45) and expressed by attenuated Arca or C
CC Salmonella typhimurium and gut-colonising bacteria for protection against
CC plague.
SQ Sequence 329 AA;

Query Match 46.6%; Score 243; DB 1; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.2e-225;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHNDLQNGIKRVEKEFLSSPTQWELRAFMVMEFSLADRIDDDILKVIYDSM 338
DB 87 ILKGGHNDLQNGIKRVEKEFLSSPTQWELRAFMVMEFSLADRIDDDILKVIYDSM 146
QY 339 NHHGDARSKLREELAEIETAEIKIYVQAEINRKLSSSGTINIHDKSINLMDKNLYGYTD 398
DB 147 NHHGDARSKLREELAEIETAEIKIYVQAEINRKLSSSGTINIHDKSINLMDKNLYGYTD 206
QY 399 EEIRKASAEYKILEKMPQTTIOVDGSEKKIYVQAEINRKLSSSGTINIHDKSINLMDKNLYGYTD 458
DB 207 EEIRKASAEYKILEKMPQTTIOVDGSEKKIYVQAEINRKLSSSGTINIHDKSINLMDKNLYGYTD 266
QY 459 NELSHFATTCSDKSRPLNDLVSOIKTQSLDITSRPNNAIEALNRFIOKYSVMORLLDDT 518
DB 267 NELSHFATTCSDKSRPLNDLVSOIKTQSLDITSRPNNAIEALNRFIOKYSVMORLLDDT 326
QY 519 SGK 521
DB 327 SGK 329

RESULT 4
ID W01045
ID W01045 standard; Protein: 480 AA.
AC W01045;
DE 28-DEC-1996 (first entry)
DE Y. pestis F1/V antigen fusion.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW F1 antigen; cafi.
OS Chimeric Yersinia pestis strain GB.
OS Chimeric synthetic.
FT Key
FT Location/Qualifiers
FT protein 1..149

FT /label="F1 antigen
FT /note="mature F1 antigen"
FT peptide 150..155
FT /label="linker
FT /note="6-amino acid peptide linker"
FT protein 156..480
FT /label="V antigen
FT /label="mature V antigen"
FT /note="mature V antigen"

PN W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI; 96-433824/43.
DR N-PSDB; T38256.
PT Yersinia pestis V antigen and F1 antigen or their protective
PS epitopic parts - useful in vaccine for protection against plague
PS Disclosure; Page 51-55; 98pp; English.
CC A fusion protein (W01045) comprises the F1 antigen (see also
CC W01042) and V antigen (see also W01041) of Yersinia pestis joined
CC by a linker that allows each protein to attain its conformational
CC state. It is the product of a gene fusion (T38256) obt'd. by PCR
CC amplification of Y. pestis DNA. F1/V fusion protein can be
CC expressed by gut-colonising organism transformants, to induce
CC an immune response against Y. pestis, the causative organism of
CC plague.
SQ Sequence 480 AA;

Query Match 39.7%; Score 207; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 6.7e-191;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHNDLQNGIKRVEKEFLSSPTQWELRAFMVMEFSLADRIDDDILKVIYDSM 338
DB 238 ILKGGHNDLQNGIKRVEKEFLSSPTQWELRAFMVMEFSLADRIDDDILKVIYDSM 297
QY 339 NHHGDARSKLREELAEIETAEIKIYVQAEINRKLSSSGTINIHDKSINLMDKNLYGYTD 398
DB 298 NHHGDARSKLREELAEIETAEIKIYVQAEINRKLSSSGTINIHDKSINLMDKNLYGYTD 357
QY 399 EEIRKASAEYKILEKMPQTTIOVDGSEKKIYVQAEINRKLSSSGTINIHDKSINLMDKNLYGYTD 458
DB 358 EEIRKASAEYKILEKMPQTTIOVDGSEKKIYVQAEINRKLSSSGTINIHDKSINLMDKNLYGYTD 417
QY 459 NELSHFATTCSDKSRPLNDLVSOIKTQ 485
DB 418 NELSHFATTCSDKSRPLNDLVSOIKTQ 444

RESULT 5
ID R76528
ID R76528 standard; Protein: 170 AA.
AC R76528;
DE 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhimurium;
KW bubonic plague; pneumonic plague.
OS Yersinia pestis.
PN W09518231-A1.
PD 06-JUL-1995.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA) UK SEC FOR DEFENCE.
PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI; 95-246396/32.
DR N-PSDB; Q92819.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response.
PT against Yersinia pestis, at mucosal surfaces.

PS Disclosure: Page 20; 27pp; English.
 CC The sequence represents the Y. pestis cafi (F1) antigen expressed
 CC from plasmid pF0R1b. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated salmonella typhimurium or salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 170 AA;

Query Match 32.6%; Score 170; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. NO. 1e-155;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 83
 DB 1 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 60
 QY 84 ELLVGLTLGGYKGTGTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 143
 DB 61 ELLVGLTLGGYKGTGTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 120
 QY 144 PVRNENLVGDVVLATGSDOFFVRSISGKGLAAGKYTDAVTVYSNQ 193
 DB 121 PVRNENLVGDVVLATGSDOFFVRSISGKGLAAGKYTDAVTVYSNQ 170

RESULT 6
 ID W01043 standard; Protein: 170 AA.
 AC W01043;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1 antigen (including signal peptide);
 KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;
 OS Yersinia pestis strain GB.
 FH Key Location/Qualifiers
 FT peptide 1..21
 FT /label- Sig_peptide

PN W0628551-A1.
 PD 19-SEP-1996.
 PF 13-MAR-1996: G00571.
 PR 13-MAR-1995: GB-005059.
 PR 15-SEP-1995: GB-018946.
 PR 05-DEC-1995: GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
 DR N-PSDB; T38248.
 PT Yersinia pestis - useful in vaccine for protection against plague
 PT epitopic parts - useful in vaccine for protection against plague
 PS Example 2; Page 61-62; 98pp; English.
 CC The F1 antigen (W01043), including the signal peptide, of Yersinia
 CC pestis was produced from a DNA sequence (T38248) obtd. by PCR
 CC amplification (see also T38257-58) of Y. pestis DNA. Expression
 CC of the F1 antigen (see also W01042) by gut-colonising organisms in
 CC the form of live vaccines can be used to protect an animal,
 CC including humans, against plague.
 SQ Sequence 170 AA;

Query Match 32.6%; Score 170; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. NO. 1e-155;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 83
 DB 1 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 60
 QY 84 ELLVGLTLGGYKGTGTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 143

DB 61 ELLVGLTLGGYKGTGTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 120
 QY 144 PVRNENLVGDVVLATGSDOFFVRSISGKGLAAGKYTDAVTVYSNQ 193
 DB 121 PVRNENLVGDVVLATGSDOFFVRSISGKGLAAGKYTDAVTVYSNQ 170

RESULT 7
 ID W59782 standard; Protein: 170 AA.
 AC W59782;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYpF1sec170.
 KW F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN W06824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-333331/29.
 DR N-PSDB; V41594.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 52; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 170 AA;

Query Match 32.6%; Score 170; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. NO. 1e-155;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 83
 DB 1 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 60
 QY 84 ELLVGLTLGGYKGTGTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 143
 DB 61 ELLVGLTLGGYKGTGTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 120
 QY 144 PVRNENLVGDVVLATGSDOFFVRSISGKGLAAGKYTDAVTVYSNQ 193
 DB 121 PVRNENLVGDVVLATGSDOFFVRSISGKGLAAGKYTDAVTVYSNQ 170

RESULT 8
 ID W59783 standard; Protein: 170 AA.
 AC W59783;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen nYpF1sec10.
 KW F1 antigen; plasmid; vaccine; plague; ds.
 OS Yersinia pestis.
 FH Key Location/Qualifiers
 FT CDS 17..532
 FT /*tag= a
 FT /*product= "F1 antigen"

PN W06824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-333331/29.
 DR N-PSDB; V41596.
 PT Protection of animals against plague - using nucleic acid encoding

PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Disclosure: Pages 53-54; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 170 AA;

Query Match 32.6%; Score 170; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1e-153;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 MKRISVIAALFGTITATNAADLTASTATATLVEPARITLYEGAPITMDNGNIDT 83
 DB 1 MKRISVIAALFGTITATNAADLTASTATATLVEPARITLYEGAPITMDNGNIDT 60
 OY 84 ELLVGTLLGKYKGTSTSVNFTDAAGDPMYLTFTSDGNNHQTFTYIGKDSRDPDIS 143
 DB 61 ELLVGTLLGKYKGTSTSVNFTDAAGDPMYLTFTSDGNNHQTFTYIGKDSRDPDIS 120
 OY 144 PVRNGENLVGDVYVATGSDPFVRSIGSKGKLAGKTTDAVTAVTSNQ 193
 DB 121 PVRNGENLVGDVYVATGSDPFVRSIGSKGKLAGKTTDAVTAVTSNQ 170

RESULT 9

ID W01044 standard; Protein; 501 AA.

AC W01044:
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1/V antigen fusion.
 KW Plaque; vaccine; genetic immunisation; V antigen; lcrV;
 KW F1 antigen; cafi.
 OS Chimeric Yersinia pestis strain GB;
 OS Chimeric synthetic.
 FH Key location/Qualifiers
 FT 1. 21
 FT /label- sig.peptide
 FT /note- "F1 antigen signal peptide"
 FT 22. 170
 FT /label- F1 antigen
 FT /note- "mature F1 antigen"
 FT 171. 176
 FT /label- linker
 FT /note- "6-amino acid peptide linker"
 FT 176. 501
 FT /label- V antigen
 FT /note- "mature V antigen"

PN W09628551-A1.
 PD 19-SEP-1996
 PF 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR N-PDB: T38249.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Example 3: Page 65-66; 98pp; English.
 CC A fusion protein (W01044) comprises the F1 antigen (see also
 CC W01043) and V antigen (see also W01041) of Yersinia pestis joined
 CC by a linker that allows each protein to attain its conformational
 CC state. It is the product of a gene fusion (T38249) obtd. by PCR
 CC amplification of Y. pestis DNA. F1/V fusion protein can be
 CC expressed by gut-colonising organism transformants, to induce
 CC an immune response against Y. pestis, the causative organism of
 CC plague.
 SQ Sequence 501 AA;

Query Match 32.6%; Score 167; DB 1; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.4e-152;
 Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 MKRISVIAALFGTITATNAADLTASTATATLVEPARITLYEGAPITMDNGNIDT 83
 DB 1 MKRISVIAALFGTITATNAADLTASTATATLVEPARITLYEGAPITMDNGNIDT 60
 OY 84 ELLVGTLLGKYKGTSTSVNFTDAAGDPMYLTFTSDGNNHQTFTYIGKDSRDPDIS 143
 DB 61 ELLVGTLLGKYKGTSTSVNFTDAAGDPMYLTFTSDGNNHQTFTYIGKDSRDPDIS 120
 OY 144 PVRNGENLVGDVYVATGSDPFVRSIGSKGKLAGKTTDAVTAVTSNQ 193
 DB 121 PVRNGENLVGDVYVATGSDPFVRSIGSKGKLAGKTTDAVTAVTSNQ 170

RESULT 10

ID R79962 standard; Protein; 329 AA.

AC R79962:
 DT 18-APR-1996 (first entry)
 DE Partial lcrV (V antigen) of Y. pestis.
 KW lcrV; V antigen; virulence; plaque; vaccine; epitope.
 OS Yersinia pestis.
 PN W09524475-A1.
 PD 14-SEP-1995.
 PF 06-MAR-1995; G00481.
 PR 08-MAR-1994; GB-004577.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Leary SEC, Titball RW, Williamson ED, Leary SE;
 DR WPI: 95-328268/42.
 DR N-PDB: T04223.
 PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
 PT oral or parenteral vaccines for protection against plague
 PS Claim 6; Page 15-16; 25pp; English.
 CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrV) encoding
 CC all or a protective epitopic part of the mature V protein of Yersinia
 CC pestis. The protein was expressed as a fusion protein with maltose
 CC binding protein or glutathione-S-transferase in 3 different plasmid
 CC vectors. Y. pestis is the highly virulent causative organism of plague
 CC in a wide range of animals, including man. The V antigen (lcrV) is an
 CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
 CC The V antigen is postulated to act as a virulence antigen, and
 CC transformed microorganisms contg. recombinant DNA encoding a V antigen
 CC protein/peptide are useful in vaccines to protect against plague.
 SQ Sequence 329 AA;

Query Match 32.1%; Score 167; DB 1; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.4e-152;
 Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 355 LFAELKIVSVIOAEINKHLSSGTTINIDKSTINLMDKNLYGTDEIRKASAEYLKEM 414
 DB 163 LFAELKIVSVIOAEINKHLSSGTTINIDKSTINLMDKNLYGTDEIRKASAEYLKEM 222
 OY 415 POTTIOVDGSEKRIYSIKDFLSEENKRTGALGNLANSYSYNNNNELSHFATTCSDKSRP 474
 DB 223 POTTIOVDGSEKRIYSIKDFLSEENKRTGALGNLANSYSYNNNNELSHFATTCSDKSRP 282
 OY 475 LMDLYSOKTQISDITSPRNSAIEALNRFQKYDSVMORLDDITSGK 521
 DB 283 LMDLYSOKTQISDITSPRNSAIEALNRFQKYDSVMORLDDITSGK 329

RESULT 11

ID W59788 standard; Protein; 149 AA.

AC W59788:
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen pYpF1mat149.

KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41609.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 63; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 149 AA;

Query Match 28.6%; Score 149; DB 1; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.4e-135;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 ADLSTATATATVEPARITITLYKGAFTITMDNGNIDTELLVGLTLYGKGTGTSTSV 104
 ID W59787
 AC W59787 standard; Protein; 150 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPfmat150.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41600.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 60; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 150 AA;

RESULT 12

ID W59787
 AC W59787 standard; Protein; 150 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPfmat150.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41600.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 60; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 150 AA;

Query Match 28.6%; Score 149; DB 1; Length 150;
 Best Local Similarity 100.0%; Pred. No. 1.4e-135;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 ADLSTATATATVEPARITITLYKGAFTITMDNGNIDTELLVGLTLYGKGTGTSTSV 104
 ID W59787
 AC W59787 standard; Protein; 150 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPfmat150.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41600.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 60; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 150 AA;

QY 105 NPTDAADPMTLTFTSODGNHGFTRVIGKDSRDPDISPKVNGENVGDDVYLATGSOD 164
 ID W59785
 AC W59785
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPfmat192.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41598.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Pages 56-57; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 192 AA;

RESULT 13

ID W59785
 AC W59785
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPfmat192.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41598.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Pages 56-57; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 192 AA;

Query Match 25.7%; Score 134; DB 1; Length 192;
 Best Local Similarity 100.0%; Pred. No. 4.5e-121;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAIALFETITANAAADTASTATATVEPARITITLYKGAFTITMDNGNIDT 83
 ID W59785
 AC W59785
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPfmat192.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41598.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Pages 56-57; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 192 AA;

RESULT 14

ID W59785
 AC W59785
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPfmat192.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41598.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Pages 56-57; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 192 AA;

PS Disclosure: Page 16; 27pp; English.
CC The sequence represents the Y. pestis cafi (F1) antigen expressed
CC from plasmid pGAL2a. The DNA construct can be used to transform
CC human or animal gut colonizing microorganisms, specifically
CC attenuated Salmonella typhimurium or Salmonella typhi. The
CC transformed microorganisms can be used as live/attenuated vaccines
CC which induce immune responses at mucosal surfaces. The vaccines
CC provide protection against infection with Y. pestis, and are
CC parenterally and orally active vaccines offering protection
CC against bubonic and pneumonic plague.
SQ Sequence 151 AA;

Query Match

24.6%; Score 128; DB 1; Length 151;

Best Local Similarity 100.0%; Pred. No. 2.1e-115; Mismatches 0; Indels 0; Gaps 0;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 TYKGAFTITMDNGNIDTELVGTLTGKTKGTSTSVNFTDAAGDPMTLFTSODGNN 125

DB 24 TYKGAFTITMDNGNIDTELVGTLTGKTKGTSTSVNFTDAAGDPMTLFTSODGNN 83

QY 126 HOFTTKVIGKDSRDFDISPRVNGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKXTDA 185

DB 84 HOFTTKVIGKDSRDFDISPRVNGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKXTDA 143

QY 186 VTVTVSNQ 193

DB 144 VTVTVSNQ 151

RESULT 15

R76527 R76527 standard; Protein; 151 AA.

AC R76527:

DT 17-DEC-1995 (first entry)

DE Yersinia pestis cafi (F1) antigen.

KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;

KM bubonic plague; pneumonic plague.

OS Yersinia pestis.

PN MO9518231-A1.

PD 06-JUL-1995.

PE 23-DEC-1994; G02818.

PR 24-DEC-1993; GB-026425.

PA (MINA) UK SEC FOR DEFENCE.

PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;

DR WPI: 95-246396/32.

DR N-PSDB; Q92818.

PT DNA constructs capable of transforming microorganisms - which can be

PT used as live or attenuated vaccines which induce an immune response,

PS Disclosure: Page 18; 27pp; English.

CC The sequence represents the Y. pestis cafi (F1) antigen expressed

CC from plasmid pGAL2a. The DNA construct can be used to transform

CC human or animal gut colonizing microorganisms, specifically

CC attenuated Salmonella typhimurium or Salmonella typhi. The

CC transformed microorganisms can be used as live/attenuated vaccines

CC which induce immune responses at mucosal surfaces. The vaccines

CC provide protection against infection with Y. pestis, and are

CC parenterally and orally active vaccines offering protection

CC against bubonic and pneumonic plague.

SQ Sequence 151 AA;

DB 84 HOFTTKVIGKDSRDFDISPRVNGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKXTDA 143

QY 186 VTVTVSNQ 193

DB 144 VTVTVSNQ 151

Search completed: August 22, 2000, 17:41:58
Job time: 2378 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:23:40 ; Search time 22.14 Seconds

(Without alignments)
360.727 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGNHHHHHHSSGHIDD.....RFLQKXDSVMRLDPTSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 152396 seqs, 15329161 residues

Word size : 0

Total number of hits satisfying chosen parameters: 152396

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_AA:*
1: /cgn2-6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2-6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2-6/ptodata/1/1aa/6.COMB.pep:*
4: /cgn2-6/ptodata/1/1aa/PCrUS.COMB.pep:*
5: /cgn2-6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243	46.6	329	2	US-08-913-477-2
2	243	46.6	329	2	US-08-913-477-4
3	243	46.6	480	2	US-08-913-477-17
4	243	46.6	501	2	US-08-913-477-23
5	131	25.1	170	2	US-08-913-477-21
6	128	24.6	151	2	US-08-913-477-11
7	24	4.6	24	3	US-08-780-496-8
8	17	3.3	21	2	US-08-651-818A-21
9	17	3.3	23	1	US-08-480-604A-24
10	17	3.3	23	2	US-08-405-496A-24
11	17	3.3	53	2	US-08-651-818A-19
12	17	3.3	54	2	US-08-651-818A-23
13	17	3.3	323	1	US-08-591-989-5
14	17	3.3	462	1	US-08-480-604A-26
15	17	3.3	462	2	US-08-405-496A-26
16	16	3.1	254	3	US-09-167-434-7
17	12	2.3	17	2	US-08-600-783-13
18	11	2.1	302	2	US-08-203-532F-4
19	11	2.1	302	4	PCT-US95-01882A-4
20	11	2.1	303	2	US-08-203-532F-2
21	11	2.1	303	4	PCT-US95-01882A-2
22	11	2.1	414	4	PCT-US92-06840-2
23	10	1.9	449	2	US-08-927-394-2
24	10	1.9	480	1	US-07-883-292-2
25	10	1.9	480	2	US-08-331-644-2
26	10	1.9	480	2	PCT-US93-04102-2
27	10	1.9	763	2	US-08-677-862-2
28	10	1.9	763	2	US-09-252-571-2

29	10	1.9	2265	2	US-08-149-097D-36	Sequence 36, Appl
30	10	1.9	2509	2	US-08-149-097D-35	Sequence 35, Appl
31	9	1.7	17	3	US-08-946-475-10	Sequence 10, Appl
32	9	1.7	20	1	US-08-373-134D-3	Sequence 3, Appl
33	9	1.7	20	2	US-09-114-637-3	Sequence 3, Appl
34	9	1.7	20	2	US-08-485-942A-99	Sequence 99, Appl
35	9	1.7	21	2	US-08-432-871C-46	Sequence 46, Appl
36	9	1.7	21	2	US-08-712-878-6	Sequence 6, Appl
37	9	1.7	21	3	US-08-485-942A-98	Sequence 98, Appl
38	9	1.7	43	2	US-08-347-563A-12	Sequence 12, Appl
39	9	1.7	43	2	US-08-485-942A-12	Sequence 12, Appl
40	9	1.7	110	1	US-08-434-705B-15	Sequence 15, Appl
41	9	1.7	110	2	US-09-086-201-15	Sequence 15, Appl
42	9	1.7	408	2	US-08-683-007A-2	Sequence 2, Appl
43	9	1.7	420	2	US-08-845-998-8	Sequence 8, Appl
44	9	1.7	711	2	US-08-946-475-9	Sequence 9, Appl
45	8	1.5	10	1	US-07-807-529A-73	Sequence 73, Appl
46	8	1.5	10	2	US-08-482-142-150	Sequence 150, App
47	8	1.5	10	2	US-08-478-572-150	Sequence 150, App
48	8	1.5	10	3	US-08-300-928C-88	Sequence 88, Appl
49	8	1.5	10	3	US-08-430-944D-88	Sequence 88, Appl
50	8	1.5	10	3	US-08-430-014-68	Sequence 88, Appl
51	8	1.5	10	4	PCT-US94-14073-17	Sequence 17, Appl
52	8	1.5	14	1	US-07-807-529A-76	Sequence 76, Appl
53	8	1.5	14	3	US-08-300-928C-91	Sequence 91, Appl
54	8	1.5	14	3	US-08-430-944D-91	Sequence 91, Appl
55	8	1.5	14	3	US-08-430-014-91	Sequence 91, Appl
56	8	1.5	15	2	US-08-467-603-53	Sequence 53, Appl
57	8	1.5	15	2	US-08-466-793-53	Sequence 53, Appl
58	8	1.5	15	2	US-08-491-861A-53	Sequence 53, Appl
59	8	1.5	19	1	US-08-346-293-7	Sequence 7, Appl
60	8	1.5	23	1	US-08-346-293-8	Sequence 8, Appl
61	8	1.5	27	1	US-08-346-293-12	Sequence 12, Appl
62	8	1.5	31	1	US-08-346-293-23	Sequence 23, Appl
63	8	1.5	38	1	US-08-346-293-4	Sequence 24, Appl
64	8	1.5	42	1	US-08-346-293-35	Sequence 25, Appl
65	8	1.5	60	1	US-08-255-457-1	Sequence 1, Appl
66	8	1.5	60	2	US-09-115-032-1	Sequence 1, Appl
67	8	1.5	60	4	PCT-US95-05772-1	Sequence 1, Appl
68	8	1.5	96	1	US-07-807-529A-39	Sequence 39, Appl
69	8	1.5	96	3	US-08-430-944D-103	Sequence 103, App
70	8	1.5	221	1	US-08-621-081A-19	Sequence 19, Appl
71	8	1.5	494	1	US-08-485-618-103	Sequence 103, App
72	8	1.5	494	2	US-08-605-672-103	Sequence 103, App
73	8	1.5	494	2	US-08-482-293A-103	Sequence 103, App
74	8	1.5	494	2	US-08-943-363-103	Sequence 103, App
75	7	1.3	7	1	US-08-100-744-7	Sequence 7, Appl
76	7	1.3	7	1	US-08-284-784-7	Sequence 7, Appl
77	7	1.3	7	2	US-08-854-811-7	Sequence 7, Appl
78	7	1.3	8	1	US-08-346-293-5	Sequence 5, Appl
79	7	1.3	9	1	US-08-155-171B-4	Sequence 4, Appl
80	7	1.3	9	2	US-08-435-998-4	Sequence 4, Appl
81	7	1.3	12	1	US-08-346-293-6	Sequence 6, Appl
82	7	1.3	14	2	US-08-448-418-100	Sequence 10, App
83	7	1.3	17	1	US-08-021-603A-19	Sequence 19, Appl
84	7	1.3	17	1	US-08-155-171B-37	Sequence 37, Appl
85	7	1.3	17	4	US-08-435-998-17	Sequence 17, Appl
86	7	1.3	17	4	PCT-US94-01176-19	Sequence 19, Appl
87	7	1.3	45	1	US-08-056-200-97	Sequence 97, Appl
88	7	1.3	45	2	US-08-800-644-57	Sequence 57, Appl
89	7	1.3	88	2	US-08-690-011A-11	Sequence 11, Appl
90	7	1.3	91	2	US-07-987-272A-11	Sequence 11, Appl
91	7	1.3	92	2	US-08-918-727-5	Sequence 5, Appl
92	7	1.3	92	2	US-09-051-589-1	Sequence 1, Appl
93	7	1.3	98	2	US-08-690-011A-19	Sequence 19, Appl
94	7	1.3	105	2	US-08-690-011A-15	Sequence 15, Appl
95	7	1.3	115	2	US-08-690-011A-21	Sequence 21, Appl
96	7	1.3	124	1	US-08-434-705B-16	Sequence 16, Appl
97	7	1.3	124	1	US-08-086-201-16	Sequence 16, Appl
98	7	1.3	210	1	US-08-234-783-2	Sequence 2, Appl
99	7	1.3	210	4	US-08-456-907-2	Sequence 2, Appl
100	7	1.3	210	4	PCT-US95-05523-2	Sequence 2, Appl
101	7	1.3	249	1	US-08-155-171B-32	Sequence 32, Appl

102	7	1	3	249	2	US-08-435-998-32	Sequence 32, Appl	175	6	1	2	12	4	PCT-US93-08062-48	Sequence 48, Appl
103	7	1	3	302	2	US-08-893-853-3	Sequence 3, Appl	176	6	1	2	12	4	PCT-US94-09700-4	Sequence 4, Appl
104	7	1	3	368	1	US-08-211-942-17	Sequence 17, Appl	177	6	1	2	13	1	US-08-571-643A-4	Sequence 22, Appl
105	7	1	3	331	1	US-08-589-981-2	Sequence 2, Appl	178	6	1	2	13	1	US-08-439-747A-2	Sequence 22, Appl
106	7	1	3	685	2	US-08-878-989-1	Sequence 1, Appl	179	6	1	2	13	1	US-08-482-577B-15	Sequence 15, Appl
107	7	1	3	685	3	US-09-136-282-2	Sequence 2, Appl	180	6	1	2	13	2	US-08-440-409B-22	Sequence 22, Appl
108	7	1	3	946	3	US-08-560-005-4	Sequence 4, Appl	181	6	1	2	14	1	US-08-226-264-7	Sequence 7, Appl
109	7	1	3	1489	5	5183745-2	Patent No. 5183745	182	6	1	2	14	1	US-08-226-264-19	Sequence 19, Appl
110	7	1	3	1794	5	5183745-6	Patent No. 5183745	183	6	1	2	14	1	US-08-226-264-20	Sequence 20, Appl
111	7	1	3	3729	5	US-08-804-227C-4	Sequence 6, Appl	184	6	1	2	14	2	US-08-467-603-58	Sequence 58, Appl
112	6	1	2	6	1	US-08-160-670A-6	Sequence 6, Appl	185	6	1	2	14	2	US-08-466-793-58	Sequence 58, Appl
113	6	1	2	6	1	US-08-131-365B-48	Sequence 1, Appl	186	6	1	2	14	2	US-08-491-861A-18	Sequence 58, Appl
114	6	1	2	6	1	US-08-469-486-47	Sequence 4, Appl	187	6	1	2	14	2	US-08-672-345C-40	Sequence 40, Appl
115	6	1	2	6	1	US-08-460-343B-51	Sequence 47, Appl	188	6	1	2	15	1	US-08-469-486-48	Sequence 48, Appl
116	6	1	2	6	1	US-08-289-567-1	Sequence 51, Appl	189	6	1	2	15	2	US-08-469-658-48	Sequence 48, Appl
117	6	1	2	6	1	US-08-398-028B-51	Sequence 51, Appl	190	6	1	2	16	1	US-08-218-025A-111	Sequence 11, App
118	6	1	2	6	1	US-08-660-626-4	Sequence 4, Appl	191	6	1	2	16	1	US-08-578-649-13	Sequence 13, Appl
119	6	1	2	6	1	US-08-434-705B-17	Sequence 17, Appl	192	6	1	2	16	1	US-08-226-264-8	Sequence 8, Appl
120	6	1	2	6	1	US-08-470-837-22	Sequence 22, Appl	193	6	1	2	16	1	PCT-US94-04361-56	Sequence 21, Appl
121	6	1	2	6	1	US-08-472-244-14	Sequence 14, Appl	194	6	1	2	16	4	US-08-323-445A-18	Sequence 18, Appl
122	6	1	2	6	2	US-08-504-265B-51	Sequence 51, Appl	195	6	1	2	17	1	US-08-515-903A-18	Sequence 18, Appl
123	6	1	2	6	2	US-08-623-833B-16	Sequence 16, Appl	196	6	1	2	17	1	US-09-136-421-10	Sequence 10, Appl
124	6	1	2	6	2	US-08-652-507-8	Sequence 8, Appl	197	6	1	2	17	3	PCT-US95-12840-18	Sequence 18, Appl
125	6	1	2	6	2	US-08-668-123-48	Sequence 48, Appl	198	6	1	2	17	4	5428012-6	Sequence 18, Appl
126	6	1	2	6	2	US-08-469-658-47	Sequence 47, Appl	199	6	1	2	17	5	5451506-6	Sequence 18, Appl
127	6	1	2	6	2	US-09-086-201-17	Sequence 17, Appl	200	6	1	2	17	5	US-08-145-006C-6	Sequence 6, Appl
128	6	1	2	6	2	US-08-591-196-57	Sequence 57, Appl	201	6	1	2	18	1	US-08-578-649-14	Sequence 14, Appl
129	6	1	2	6	2	US-08-596-139-18	Sequence 18, Appl	202	6	1	2	18	1	US-08-356-060A-47	Sequence 47, Appl
130	6	1	2	6	3	US-08-481-435-42	Sequence 42, Appl	203	6	1	2	18	2	US-08-670-175-8	Sequence 8, Appl
131	6	1	2	6	3	US-08-772-440-28	Sequence 28, Appl	204	6	1	2	18	2	US-08-895-707-15	Sequence 15, Appl
132	6	1	2	6	4	PCT-US94-04361-55	Sequence 55, Appl	205	6	1	2	18	3	PCT-US94-00545-6	Sequence 6, Appl
133	6	1	2	6	4	US-08-762-433-6	Sequence 6, Appl	206	6	1	2	18	4	US-08-462-436-13	Sequence 13, Appl
134	6	1	2	6	8	US-09-001-219-6	Sequence 6, Appl	207	6	1	2	19	3	US-09-167-434-15	Sequence 15, Appl
135	6	1	2	6	8	US-08-481-658B-22	Sequence 22, Appl	208	6	1	2	20	1	US-08-416-336-4	Sequence 4, Appl
136	6	1	2	6	8	US-08-477-504A-22	Sequence 22, Appl	209	6	1	2	20	1	US-08-543-020-12	Sequence 12, Appl
137	6	1	2	6	8	US-08-486-756A-22	Sequence 22, Appl	210	6	1	2	20	2	US-08-480-473B-16	Sequence 16, Appl
138	6	1	2	6	8	US-08-485-862B-22	Sequence 22, Appl	211	6	1	2	21	3	US-08-915-213-46	Sequence 46, Appl
139	6	1	2	6	8	US-08-487-739-22	Sequence 22, Appl	212	6	1	2	21	3	US-08-127-909-14	Sequence 14, Appl
140	6	1	2	6	8	US-08-816-977-11	Sequence 11, Appl	213	6	1	2	22	2	US-08-457-353-14	Sequence 14, Appl
141	6	1	2	6	8	US-08-284-784-12	Sequence 12, Appl	214	6	1	2	22	2	US-08-323-445A-18	Sequence 18, Appl
142	6	1	2	6	8	US-08-323-445A-2	Sequence 2, Appl	215	6	1	2	23	2	US-08-350-260A-42	Sequence 42, Appl
143	6	1	2	6	9	US-08-106-744-12	Sequence 12, Appl	216	6	1	2	23	2	PCT-US94-01176-16	Sequence 16, Appl
144	6	1	2	6	9	US-08-823-445A-2	Sequence 2, Appl	217	6	1	2	24	1	US-08-127-909-19	Sequence 19, Appl
145	6	1	2	6	9	US-08-515-903A-2	Sequence 12, Appl	218	6	1	2	24	4	US-08-457-353-19	Sequence 19, Appl
146	6	1	2	6	9	US-08-737-757-12	Sequence 12, Appl	219	6	1	2	25	1	US-08-473-475A-21	Sequence 21, Appl
147	6	1	2	6	9	US-08-854-811-12	Sequence 19, Appl	220	6	1	2	25	2	US-08-227-372-2	Sequence 2, Appl
148	6	1	2	6	9	US-09-143-470-19	Sequence 19, Appl	221	6	1	2	26	1	US-08-657-641-12	Sequence 12, Appl
149	6	1	2	6	9	US-08-895-707-9	Sequence 9, Appl	222	6	1	2	26	2	US-08-470-397-2	Sequence 2, Appl
150	6	1	2	6	9	PCT-US95-12840-2	Sequence 2, Appl	223	6	1	2	26	3	US-08-895-707-21	Sequence 21, Appl
151	6	1	2	6	9	US-08-459-064B-16	Sequence 16, Appl	224	6	1	2	26	3	PCT-US94-07233-12	Sequence 12, Appl
152	6	1	2	6	10	US-08-737-757-10	Sequence 10, Appl	225	6	1	2	26	4	PCT-US94-05150-37	Sequence 37, Appl
153	6	1	2	6	10	US-08-628-665-14	Sequence 14, Appl	226	6	1	2	29	4	US-08-663-227-37	Sequence 37, Appl
154	6	1	2	6	10	US-08-524-757-30	Sequence 30, Appl	227	6	1	2	31	2	US-08-323-445A-17	Sequence 17, Appl
155	6	1	2	6	10	US-08-524-757-30	Sequence 30, Appl	228	6	1	2	32	1	US-08-515-903A-17	Sequence 17, Appl
156	6	1	2	6	10	US-08-460-421A-16	Sequence 16, Appl	229	6	1	2	32	3	PCT-US95-12840-17	Sequence 17, Appl
157	6	1	2	6	10	US-08-717-169-13	Sequence 13, Appl	230	6	1	2	32	3	US-08-670-175-8	Sequence 8, Appl
158	6	1	2	6	10	US-08-320-240A-3	Sequence 3, Appl	231	6	1	2	33	4	US-08-878-865-38	Sequence 38, Appl
159	6	1	2	6	10	US-08-792-553-24	Sequence 24, Appl	232	6	1	2	36	3	US-08-273-146-39	Sequence 39, Appl
160	6	1	2	6	10	PCT-US93-00909-13	Sequence 13, Appl	233	6	1	2	39	2	US-08-457-353-19	Sequence 19, Appl
161	6	1	2	6	10	PCT-US95-04439-9	Sequence 9, Appl	234	6	1	2	40	2	US-08-664-449-39	Sequence 39, Appl
162	6	1	2	6	10	US-08-133-804-11	Sequence 11, Appl	235	6	1	2	42	1	US-08-273-146-41	Sequence 41, Appl
163	6	1	2	6	10	US-08-461-838-11	Sequence 11, Appl	236	6	1	2	43	2	US-08-652-816A-43	Sequence 43, Appl
164	6	1	2	6	11	US-08-737-757-4	Sequence 4, Appl	237	6	1	2	44	2	US-08-652-816A-43	Sequence 43, Appl
165	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	238	6	1	2	45	2	US-08-220-378-1	Sequence 1, Appl
166	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	239	6	1	2	45	2	US-08-696-012-1	Sequence 1, Appl
167	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	240	6	1	2	48	1	US-08-889-291-3	Sequence 3, Appl
168	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	241	6	1	2	55	1	US-08-457-245-21	Sequence 21, Appl
169	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	242	6	1	2	55	1	US-08-466-033-23	Sequence 23, App
170	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	243	6	1	2	55	1	US-08-444-733-273	Sequence 273, App
171	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	244	6	1	2	55	1		
172	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	245	6	1	2	55	1		
173	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	246	6	1	2	55	1		
174	6	1	2	6	11	US-08-461-838-11	Sequence 11, Appl	247	6	1	2	55	1		

248	6	1.2	55	2	US-08-464-134-273	Sequence 273, App	321	6	1.2	233	3	US-08-837-056-9	Sequence 9, App11
249	6	1.2	55	2	US-08-461-361-273	Sequence 273, App	322	6	1.2	242	3	US-09-167-434-8	Sequence 8, App11
250	6	1.2	55	2	US-08-485-910-273	Sequence 273, App	323	6	1.2	243	3	US-09-167-434-8	Sequence 9, App11
251	6	1.2	57	2	US-08-652-816A-41	Sequence 41, App1	324	6	1.2	244	3	US-08-480-753-4	Sequence 4, App11
252	6	1.2	57	2	US-08-958-201-18	Sequence 18, App1	325	6	1.2	246	3	US-09-167-434-11	Sequence 11, App11
253	6	1.2	67	2	US-08-142-551B-9	Sequence 9, App11	326	6	1.2	248	4	PCT-US93-03895-8	Sequence 8, App11
254	6	1.2	69	2	US-08-687-865A-21	Sequence 21, App1	327	6	1.2	250	4	PCT-US93-03895-6	Sequence 6, App11
255	6	1.2	71	3	US-08-783-974-6	Sequence 6, App11	328	6	1.2	251	1	US-08-397-633A-74	Sequence 74, App1
256	6	1.2	79	3	US-08-448-418-97	Sequence 97, App1	329	6	1.2	252	1	US-08-792-824-10	Sequence 10, App1
257	6	1.2	81	2	US-08-717-169-15	Sequence 15, App1	330	6	1.2	253	3	US-08-988-818-13	Sequence 13, App1
258	6	1.2	85	3	US-08-772-440-32	Sequence 32, App1	331	6	1.2	254	2	US-08-792-824-7	Sequence 7, App11
259	6	1.2	87	1	US-08-160-670A-7	Sequence 7, App1	332	6	1.2	254	2	US-08-792-824-7	Sequence 7, App11
260	6	1.2	89	1	US-08-340-812-13	Sequence 13, App1	333	6	1.2	254	2	US-08-792-824-10	Sequence 10, App1
261	6	1.2	89	1	US-08-459-064B-13	Sequence 13, App1	334	6	1.2	254	2	US-08-792-824-13	Sequence 13, App1
262	6	1.2	89	2	US-08-460-421A-13	Sequence 13, App1	335	6	1.2	255	3	US-09-167-434-10	Sequence 10, App1
263	6	1.2	89	2	US-08-717-169-7	Sequence 7, App1	336	6	1.2	257	3	US-08-486-099-113	Sequence 113, App
264	6	1.2	89	3	US-08-946-026-2	Sequence 2, App11	337	6	1.2	257	3	US-08-360-107A-123	Sequence 123, App
265	6	1.2	93	2	US-08-690-011A-9	Sequence 9, App11	338	6	1.2	257	3	US-08-484-223B-113	Sequence 113, App
266	6	1.2	94	2	US-08-266-451B-20	Sequence 20, App1	339	6	1.2	257	3	US-08-919-597-113	Sequence 113, App
267	6	1.2	94	2	US-08-748-725-20	Sequence 20, App1	340	6	1.2	257	3	US-08-919-597-113	Sequence 113, App
268	6	1.2	97	3	US-08-816-977-23	Sequence 23, App1	341	6	1.2	257	3	US-08-475-668A-113	Sequence 113, App
269	6	1.2	97	3	US-08-816-977-27	Sequence 27, App1	342	6	1.2	257	3	US-08-485-551A-113	Sequence 113, App
270	6	1.2	98	3	US-08-478-097A-4	Sequence 4, App11	343	6	1.2	261	1	US-08-160-670A-45	Sequence 45, App
271	6	1.2	99	2	US-08-672-345C-94	Sequence 94, App11	344	6	1.2	262	1	US-08-397-633A-73	Sequence 73, App1
272	6	1.2	99	2	US-08-481-905-62	Sequence 62, App1	345	6	1.2	263	2	US-08-353-476-66	Sequence 66, App1
273	6	1.2	99	3	US-08-481-985B-62	Sequence 62, App1	346	6	1.2	263	2	US-08-353-476-67	Sequence 67, App1
274	6	1.2	109	2	US-08-672-345C-4	Sequence 4, App11	347	6	1.2	264	2	US-08-436-74B-8	Sequence 8, App11
275	6	1.2	112	1	US-08-211-202-133	Sequence 133, App	348	6	1.2	264	2	US-07-857-224B-9	Sequence 9, App11
276	6	1.2	116	1	US-08-478-039-67	Sequence 67, App1	349	6	1.2	265	3	US-08-483-857-8	Sequence 8, App11
277	6	1.2	116	1	US-08-476-349A-67	Sequence 67, App1	350	6	1.2	275	2	US-08-578-709-13	Sequence 13, App1
278	6	1.2	116	1	US-08-481-377-26	Sequence 26, App1	351	6	1.2	281	1	US-08-397-633A-75	Sequence 75, App1
279	6	1.2	116	2	US-08-491-835-24	Sequence 24, App1	352	6	1.2	287	1	US-08-397-633A-76	Sequence 76, App1
280	6	1.2	116	3	US-09-153-155A-35	Sequence 35, App1	353	6	1.2	297	1	US-08-011-398B-4	Sequence 4, App11
281	6	1.2	116	4	PCT-US94-00666-26	Sequence 26, App1	354	6	1.2	297	1	US-08-464-051-4	Sequence 4, App11
282	6	1.2	116	4	PCT-US94-00666-24	Sequence 24, App1	355	6	1.2	297	2	US-08-464-496-4	Sequence 4, App11
283	6	1.2	120	4	PCT-US93-03895-2	Sequence 2, App11	356	6	1.2	297	2	US-08-554-385-3	Sequence 3, App11
284	6	1.2	121	4	PCT-US93-03895-4	Sequence 4, App11	357	6	1.2	298	3	US-09-080-044-9	Sequence 9, App11
285	6	1.2	126	1	US-08-268-348A-12	Sequence 12, App1	358	6	1.2	301	2	US-08-661-051-14	Sequence 14, App1
286	6	1.2	127	1	US-08-253-155A-35	Sequence 35, App1	359	6	1.2	304	1	US-08-132-405-4	Sequence 4, App11
287	6	1.2	128	1	US-08-426-627-14	Sequence 14, App1	360	6	1.2	304	1	US-08-393-938A-4	Sequence 4, App11
288	6	1.2	133	1	US-08-268-348A-8	Sequence 8, App11	361	6	1.2	304	4	PCT-US91-01861-4	Sequence 4, App11
289	6	1.2	133	1	US-08-268-348A-10	Sequence 10, App1	362	6	1.2	308	2	US-08-793-958-2	Sequence 2, App11
290	6	1.2	172	2	US-08-557-309B-39	Sequence 39, App1	363	6	1.2	308	2	US-08-499-568-2	Sequence 2, App11
291	6	1.2	172	3	US-08-834-306-39	Sequence 39, App1	364	6	1.2	311	3	US-08-946-914-4	Sequence 4, App11
292	6	1.2	174	3	US-08-906-136A-2	Sequence 2, App11	365	6	1.2	312	1	US-08-414-625-2	Sequence 2, App11
293	6	1.2	174	3	US-08-806-136A-4	Sequence 4, App11	366	6	1.2	315	2	US-08-484-956-91	Sequence 91, App1
294	6	1.2	179	2	US-08-809-267-7	Sequence 7, App11	367	6	1.2	315	2	US-08-757-653-91	Sequence 91, App1
295	6	1.2	179	4	PCT-US95-13662A-7	Sequence 7, App11	368	6	1.2	319	3	US-08-630-172-12	Sequence 12, App1
296	6	1.2	180	3	US-08-772-440-31	Sequence 31, App1	369	6	1.2	320	2	US-08-757-653-153	Sequence 153, App
297	6	1.2	183	5	5168049-3	Patent No. 5168049	370	6	1.2	320	2	US-08-823-516-61	Sequence 61, App1
298	6	1.2	184	3	US-09-023-082A-30	Sequence 30, App1	371	6	1.2	323	3	US-08-816-977-21	Sequence 21, App1
299	6	1.2	197	2	US-08-879-561-5	Sequence 5, App11	372	6	1.2	326	3	US-08-816-977-25	Sequence 25, App1
300	6	1.2	198	1	US-08-642-255-131	Sequence 131, App	373	6	1.2	326	3	US-08-816-977-25	Sequence 25, App1
301	6	1.2	198	1	US-08-397-633A-41	Sequence 41, App1	374	6	1.2	343	2	US-09-123-851-1	Sequence 1, App11
302	6	1.2	198	1	US-08-397-633A-52	Sequence 52, App1	375	6	1.2	343	2	US-08-447-430A-43	Sequence 43, App1
303	6	1.2	214	2	US-08-846-790A-1	Sequence 1, App11	376	6	1.2	344	1	US-07-941-552-24	Sequence 24, App1
304	6	1.2	214	2	US-08-935-333-1	Sequence 1, App11	377	6	1.2	351	2	US-08-701-191A-3	Sequence 3, App11
305	6	1.2	218	1	US-08-644-664B-14	Sequence 14, App1	378	6	1.2	357	1	US-08-078-683A-8	Sequence 8, App11
306	6	1.2	218	2	US-08-761-277-14	Sequence 14, App1	379	6	1.2	360	1	US-08-674-612-3	Sequence 3, App11
307	6	1.2	219	1	US-08-621-081A-16	Sequence 16, App1	380	6	1.2	360	1	US-08-469-421-14	Sequence 14, App1
308	6	1.2	223	1	US-08-430-633-1	Sequence 1, App11	381	6	1.2	360	1	US-08-250-975-13	Sequence 13, App1
309	6	1.2	223	1	US-08-336-854-1	Sequence 1, App11	382	6	1.2	360	1	US-08-920-296-3	Sequence 3, App11
310	6	1.2	224	1	US-08-360-673-7	Sequence 7, App11	383	6	1.2	360	2	US-08-605-002A-14	Sequence 14, App1
311	6	1.2	224	1	US-08-414-625-4	Sequence 4, App11	384	6	1.2	360	2	US-08-950-448A-14	Sequence 14, App1
312	6	1.2	227	5	5428012-3	Patent No. 5428012	385	6	1.2	360	2	US-08-554-385-20	Sequence 20, App1
313	6	1.2	227	5	5428012-4	Patent No. 5428012	386	6	1.2	360	4	PCT-US94-10529-14	Sequence 14, App1
314	6	1.2	227	5	5451506-3	Patent No. 5451506	387	6	1.2	376	3	US-08-751-513-8	Sequence 8, App11
315	6	1.2	228	5	5428012-2	Patent No. 5428012	388	6	1.2	377	2	US-08-929-922B-2	Sequence 2, App11
316	6	1.2	228	5	5451506-2	Patent No. 5451506	389	6	1.2	378	2	US-08-707-237A-104	Sequence 104, App
317	6	1.2	233	1	US-08-446-918A-4	Sequence 4, App11	390	6	1.2	378	2	US-08-642-246-26	Sequence 26, App1
318	6	1.2	233	2	US-08-480-753-2	Sequence 2, App11	391	6	1.2	380	4	PCT-US96-06229-26	Sequence 26, App1
319	6	1.2	233	2	US-08-80-806-4	Sequence 4, App11	392	6	1.2	380	1	US-08-585-758A-4	Sequence 4, App11
320	6	1.2	233	3	US-09-041-889-9	Sequence 9, App11	393	6	1.2	380	1	US-08-977-818-4	Sequence 4, App11

394	6	1.2	380	2	US-08-670-274B-4	Sequence 4, Appl	467	6	1.2	593	1	US-08-374-155A-14	Sequence 14, Appl
395	6	1.2	381	1	US-08-585-758A-2	Sequence 2, Appl	468	6	1.2	593	2	US-08-785-396-14	Sequence 14, Appl
396	6	1.2	381	1	US-08-977-818-2	Sequence 2, Appl	469	6	1.2	601	2	US-08-606-288-7	Sequence 7, Appl
397	6	1.2	381	2	US-08-670-274B-2	Sequence 2, Appl	470	6	1.2	601	2	US-08-606-288-10	Sequence 10, Appl
398	6	1.2	381	2	US-08-786-999-3	Sequence 3, Appl	471	6	1.2	612	2	US-08-752-307B-11	Sequence 11, Appl
399	6	1.2	386	3	US-08-895-707-7	Sequence 7, Appl	472	6	1.2	616	3	US-08-895-707-2	Sequence 2, Appl
400	6	1.2	390	2	US-08-786-999-1	Sequence 1, Appl	473	6	1.2	628	2	US-08-394-177-6	Sequence 6, Appl
401	6	1.2	391	1	US-07-995-657-6	Sequence 6, Appl	474	6	1.2	630	2	US-08-596-319-2	Sequence 2, Appl
402	6	1.2	391	1	US-08-474-587-6	Sequence 6, Appl	475	6	1.2	631	1	US-08-605-541B-12	Sequence 12, Appl
403	6	1.2	391	2	US-08-244-951A-10	Sequence 10, Appl	476	6	1.2	662	2	US-08-770-761A-5	Sequence 5, Appl
404	6	1.2	391	2	US-08-389-011-73	Sequence 23, Appl	477	6	1.2	677	3	US-08-480-640A-115	Sequence 115, App
405	6	1.2	391	3	US-08-403-917A-23	Sequence 23, Appl	478	6	1.2	677	3	US-08-480-640A-193	Sequence 193, App
406	6	1.2	391	4	PCT-US95-15696-2	Sequence 2, Appl	479	6	1.2	680	1	US-08-542-363-4	Sequence 4, Appl
407	6	1.2	393	4	US-08-977-554-2	Sequence 2, Appl	480	6	1.2	682	1	US-08-642-255-126	Sequence 126, App
408	6	1.2	401	5	5252556-1	Patent No. 5252556	481	6	1.2	682	1	US-08-397-633A-36	Sequence 36, Appl
409	6	1.2	403	2	US-09-061-337-10	Sequence 10, Appl	482	6	1.2	682	1	US-08-481-635-6	Sequence 6, Appl
410	6	1.2	403	2	US-09-123-139-10	Sequence 10, Appl	483	6	1.2	701	2	US-08-533-669A-16	Sequence 16, Appl
411	6	1.2	403	2	US-09-340-991-10	Sequence 10, Appl	484	6	1.2	705	2	US-08-770-761A-7	Sequence 7, Appl
412	6	1.2	406	2	US-08-222-719-1	Sequence 1, Appl	485	6	1.2	729	1	US-07-640-029-3	Sequence 3, Appl
413	6	1.2	406	2	US-08-470-925-1	Sequence 1, Appl	486	6	1.2	731	1	US-07-921-807B-5	Sequence 5, Appl
414	6	1.2	406	2	US-08-471-613-1	Sequence 1, Appl	487	6	1.2	731	1	US-08-441-944A-5	Sequence 5, Appl
415	6	1.2	406	2	PCT-US93-10443-1	Sequence 1, Appl	488	6	1.2	733	1	US-07-640-029-4	Sequence 4, Appl
416	6	1.2	407	2	US-08-776-585-3	Sequence 3, Appl	489	6	1.2	733	1	US-08-441-944A-16	Sequence 6, Appl
417	6	1.2	423	2	US-08-910-798-2	Sequence 2, Appl	490	6	1.2	733	1	US-08-188-582-16	Sequence 16, Appl
418	6	1.2	423	2	US-08-313-533-13	Sequence 13, Appl	491	6	1.2	737	1	US-08-646-715-16	Sequence 16, Appl
419	6	1.2	425	3	US-08-767-993-13	Sequence 13, Appl	492	6	1.2	741	2	US-08-849-480A-9	Sequence 9, Appl
420	6	1.2	427	2	US-08-846-021A-8	Sequence 8, Appl	493	6	1.2	761	2	US-08-707-337A-84	Sequence 84, Appl
421	6	1.2	439	1	US-08-507-431-4	Sequence 2, Appl	494	6	1.2	762	1	US-08-642-255-114	Sequence 114, App
422	6	1.2	439	2	US-08-700-546-2	Sequence 2, Appl	495	6	1.2	762	1	US-08-642-255-120	Sequence 120, App
423	6	1.2	439	3	US-08-902-655A-4	Sequence 4, Appl	496	6	1.2	762	1	US-08-397-633A-26	Sequence 26, App
424	6	1.2	439	3	US-09-116-622-4	Sequence 4, Appl	497	6	1.2	762	1	US-08-397-633A-126	Sequence 126, App
425	6	1.2	450	2	US-08-665-037-2	Sequence 2, Appl	498	6	1.2	764	1	US-08-424-567-2	Sequence 2, Appl
426	6	1.2	450	2	US-08-666-067-2	Sequence 2, Appl	499	6	1.2	764	2	US-08-424-567-2	Sequence 2, Appl
427	6	1.2	450	2	US-08-733-870-2	Sequence 2, Appl	500	6	1.2	768	2	US-08-560-398-2	Sequence 2, Appl
428	6	1.2	458	5	5217891-15	Patent No. 5217891	501	6	1.2	770	1	US-08-369-796-12	Sequence 12, Appl
429	6	1.2	474	2	US-09-096-982-8	Sequence 8, Appl	502	6	1.2	770	1	US-08-416-581B-1	Sequence 1, Appl
430	6	1.2	474	2	US-08-635-650A-8	Sequence 8, Appl	503	6	1.2	770	1	US-08-416-581B-5	Sequence 5, Appl
431	6	1.2	479	1	US-08-397-633A-78	Sequence 78, Appl	504	6	1.2	770	1	US-08-416-581B-9	Sequence 9, Appl
432	6	1.2	481	2	US-08-477-451-19	Sequence 19, Appl	505	6	1.2	770	2	US-08-852-091-12	Sequence 12, Appl
433	6	1.2	486	1	US-08-397-633A-77	Sequence 77, Appl	506	6	1.2	770	2	US-08-852-091-12	Sequence 12, Appl
434	6	1.2	494	4	US-08-464-340A-4	Sequence 4, Appl	507	6	1.2	770	2	US-08-820-754-12	Sequence 12, Appl
435	6	1.2	494	4	PCT-US94-08449A-4	Sequence 4, Appl	508	6	1.2	770	3	US-08-956-652-12	Sequence 12, Appl
436	6	1.2	498	2	US-07-916-098A-4	Sequence 4, Appl	509	6	1.2	770	4	PCT-US95-17025-12	Sequence 12, Appl
437	6	1.2	500	2	US-08-578-709-15	Sequence 15, Appl	510	6	1.2	771	1	US-08-276-099A-14	Sequence 14, Appl
438	6	1.2	504	2	US-08-896-005-3	Sequence 3, Appl	511	6	1.2	771	1	US-08-781-690-14	Sequence 14, Appl
439	6	1.2	511	3	US-09-105-039A-2	Sequence 2, Appl	512	6	1.2	777	2	US-08-477-396A-4	Sequence 4, Appl
440	6	1.2	514	2	US-08-865-311-2	Sequence 2, Appl	513	6	1.2	779	1	US-08-426-627-24	Sequence 24, Appl
441	6	1.2	519	1	US-08-589-446-8	Sequence 8, Appl	514	6	1.2	779	1	US-08-426-627-24	Sequence 24, Appl
442	6	1.2	519	1	US-08-444-882-8	Sequence 8, Appl	515	6	1.2	780	1	US-08-375-709-13	Sequence 13, Appl
443	6	1.2	519	2	US-08-389-459A-8	Sequence 8, Appl	516	6	1.2	780	1	US-08-752-929-13	Sequence 13, Appl
444	6	1.2	519	3	US-08-987-867A-8	Sequence 8, Appl	517	6	1.2	788	1	US-08-194-338-12	Sequence 12, Appl
445	6	1.2	524	1	US-08-447-500-24	Sequence 24, Appl	518	6	1.2	788	1	US-08-572-225-1	Sequence 1, Appl
446	6	1.2	524	1	US-08-454-097-24	Sequence 24, Appl	519	6	1.2	811	1	US-08-426-627-2	Sequence 2, Appl
447	6	1.2	524	3	US-08-453-866-24	Sequence 24, Appl	520	6	1.2	811	1	US-08-426-627-22	Sequence 22, Appl
448	6	1.2	524	3	US-08-185-359-24	Sequence 24, Appl	521	6	1.2	816	1	US-07-640-029-1	Sequence 1, Appl
449	6	1.2	525	1	US-08-375-709-19	Sequence 19, Appl	522	6	1.2	817	1	US-07-640-029-2	Sequence 2, Appl
450	6	1.2	525	1	US-08-752-929-19	Sequence 19, Appl	523	6	1.2	820	1	US-07-921-807B-3	Sequence 3, Appl
451	6	1.2	526	2	US-08-852-401-3	Sequence 3, Appl	524	6	1.2	820	1	US-08-291-996-2	Sequence 2, Appl
452	6	1.2	527	1	US-08-160-670A-8	Sequence 8, Appl	525	6	1.2	820	1	US-08-441-944A-3	Sequence 3, Appl
453	6	1.2	530	3	US-09-105-039A-4	Sequence 4, Appl	526	6	1.2	820	1	US-08-166-717D-6	Sequence 6, Appl
454	6	1.2	532	3	US-08-737-336-6	Sequence 6, Appl	527	6	1.2	820	2	US-08-485-278-2	Sequence 2, Appl
455	6	1.2	537	3	US-08-920-610-2	Sequence 2, Appl	528	6	1.2	822	1	US-07-997-133-1	Sequence 1, Appl
456	6	1.2	539	3	US-08-906-360-1	Sequence 1, Appl	529	6	1.2	822	1	US-07-921-807B-4	Sequence 4, Appl
457	6	1.2	547	2	US-08-785-048-3	Sequence 3, Appl	530	6	1.2	822	1	US-08-459-296-2	Sequence 2, Appl
458	6	1.2	547	2	US-08-966-799-3	Sequence 3, Appl	531	6	1.2	822	2	US-08-441-944A-4	Sequence 4, Appl
459	6	1.2	550	4	PCT-US93-03027-1	Sequence 1, Appl	532	6	1.2	822	2	US-08-451-822A-12	Sequence 12, Appl
460	6	1.2	553	1	US-08-785-048-2	Sequence 2, Appl	533	6	1.2	825	1	US-08-642-255-132	Sequence 132, App
461	6	1.2	553	2	US-08-661-052-16	Sequence 16, Appl	534	6	1.2	829	1	US-08-337-633A-53	Sequence 53, Appl
462	6	1.2	553	2	US-08-996-799-2	Sequence 2, Appl	535	6	1.2	832	3	US-08-630-820-7	Sequence 7, Appl
463	6	1.2	559	1	US-08-320-559-31	Sequence 31, Appl	536	6	1.2	836	1	US-08-426-627-6	Sequence 6, Appl
464	6	1.2	559	1	US-08-545-860D-31	Sequence 31, Appl	537	6	1.2	837	1	US-08-426-627-23	Sequence 23, Appl
465	6	1.2	559	4	PCT-US94-04496-31	Sequence 31, Appl	538	6	1.2	856	2	US-07-916-098A-2	Sequence 2, Appl
466	6	1.2	561	1	US-08-360-673-2	Sequence 2, Appl	539	6	1.2	870	2	US-08-732-192A-2	Sequence 2, Appl

248	6	1.2	55	2	US-08-464-134-273	Sequence 273, App	321	6	1.2	233	3	US-08-837-058-9	Sequence 9, Appl
249	6	1.2	55	2	US-08-461-361-273	Sequence 273, App	322	6	1.2	242	3	US-09-167-434-8	Sequence 8, Appl
250	6	1.2	55	2	US-08-485-910-273	Sequence 273, App	323	6	1.2	243	3	US-09-167-434-9	Sequence 9, Appl
251	6	1.2	57	2	US-08-652-816A-41	Sequence 41, Appl	324	6	1.2	244	2	US-08-480-753-4	Sequence 4, Appl
252	6	1.2	57	2	US-08-958-201-18	Sequence 18, Appl	325	6	1.2	246	3	US-09-167-434-11	Sequence 11, Appl
253	6	1.2	67	2	US-08-142-551B-9	Sequence 9, Appl	326	6	1.2	248	4	PCT-US93-03895-6	Sequence 6, Appl
254	6	1.2	69	2	US-08-687-865A-21	Sequence 21, Appl	327	6	1.2	250	4	PCT-US93-03895-8	Sequence 8, Appl
255	6	1.2	71	3	US-08-783-974-6	Sequence 6, Appl	328	6	1.2	251	1	US-08-397-633A-74	Sequence 74, Appl
256	6	1.2	79	2	US-08-448-418-57	Sequence 97, Appl	329	6	1.2	252	1	US-08-792-019B-13	Sequence 13, Appl
257	6	1.2	81	2	US-08-717-159-15	Sequence 15, Appl	330	6	1.2	252	3	US-08-988-819-13	Sequence 13, Appl
258	6	1.2	85	3	US-08-712-440-32	Sequence 32, Appl	331	6	1.2	254	2	US-08-792-824-4	Sequence 4, Appl
259	6	1.2	87	1	US-08-160-670A-7	Sequence 7, Appl	332	6	1.2	254	2	US-08-792-824-10	Sequence 10, Appl
260	6	1.2	89	1	US-08-340-812-13	Sequence 13, Appl	333	6	1.2	254	2	US-08-792-824-13	Sequence 13, Appl
261	6	1.2	89	1	US-08-459-064B-13	Sequence 13, Appl	334	6	1.2	255	3	US-09-167-434-13	Sequence 13, Appl
262	6	1.2	89	2	US-08-460-421A-13	Sequence 13, Appl	335	6	1.2	255	3	US-08-486-099-113	Sequence 113, App
263	6	1.2	89	2	US-08-717-159-7	Sequence 7, Appl	336	6	1.2	257	3	US-08-360-107A-113	Sequence 113, App
264	6	1.2	89	3	US-08-946-026-2	Sequence 2, Appl	337	6	1.2	257	3	US-08-484-223B-113	Sequence 113, App
265	6	1.2	93	2	US-08-690-011A-9	Sequence 9, Appl	338	6	1.2	257	3	US-08-919-597-113	Sequence 113, App
266	6	1.2	94	1	US-08-266-451B-20	Sequence 20, Appl	339	6	1.2	257	3	US-08-475-668A-113	Sequence 113, App
267	6	1.2	94	2	US-08-748-725-50	Sequence 23, Appl	340	6	1.2	257	3	US-08-485-551A-113	Sequence 113, App
268	6	1.2	97	3	US-08-816-977-23	Sequence 27, Appl	341	6	1.2	257	3	US-08-160-670A-45	Sequence 45, App
269	6	1.2	97	3	US-08-816-977-27	Sequence 27, Appl	342	6	1.2	261	1	US-08-397-633A-73	Sequence 73, Appl
270	6	1.2	98	3	US-08-478-097A-4	Sequence 4, Appl	343	6	1.2	262	1	US-08-160-670A-32	Sequence 32, Appl
271	6	1.2	99	2	US-08-672-345C-94	Sequence 94, Appl	344	6	1.2	263	2	US-08-353-476-66	Sequence 66, Appl
272	6	1.2	99	2	US-08-484-905-62	Sequence 62, Appl	345	6	1.2	263	2	US-08-436-748-8	Sequence 8, Appl
273	6	1.2	99	3	US-08-481-985B-62	Sequence 62, Appl	346	6	1.2	264	2	US-07-857-224B-9	Sequence 9, Appl
274	6	1.2	109	2	US-08-672-345C-4	Sequence 4, Appl	347	6	1.2	264	2	US-08-483-857-8	Sequence 8, Appl
275	6	1.2	112	1	US-08-211-202-133	Sequence 133, App	348	6	1.2	265	3	US-08-578-709-13	Sequence 13, Appl
276	6	1.2	116	1	US-08-478-039-67	Sequence 67, Appl	349	6	1.2	275	2	US-08-397-633A-75	Sequence 75, Appl
277	6	1.2	116	1	US-08-476-349A-67	Sequence 67, Appl	350	6	1.2	281	1	US-08-397-633A-76	Sequence 76, Appl
278	6	1.2	116	1	US-08-481-377-26	Sequence 26, Appl	351	6	1.2	287	1	US-08-011-398B-4	Sequence 4, Appl
279	6	1.2	116	2	US-08-491-835-24	Sequence 24, Appl	352	6	1.2	297	1	US-08-464-051-4	Sequence 4, Appl
280	6	1.2	116	3	US-09-153-733A-26	Sequence 26, Appl	353	6	1.2	297	2	US-08-462-198-4	Sequence 4, Appl
281	6	1.2	116	3	PCT-US94-00666-26	Sequence 26, Appl	354	6	1.2	297	2	US-08-554-485-3	Sequence 3, Appl
282	6	1.2	116	4	PCT-US94-00685-24	Sequence 24, Appl	355	6	1.2	297	2	US-08-499-568-2	Sequence 2, Appl
283	6	1.2	120	4	PCT-US93-03895-2	Sequence 2, Appl	356	6	1.2	301	2	US-08-132-405-4	Sequence 4, Appl
284	6	1.2	121	4	PCT-US93-03895-4	Sequence 4, Appl	357	6	1.2	304	1	US-08-395-939A-4	Sequence 4, Appl
285	6	1.2	125	1	US-08-268-348A-12	Sequence 12, Appl	358	6	1.2	304	1	PCT-US91-01861-4	Sequence 4, Appl
286	6	1.2	127	1	US-08-253-155A-35	Sequence 35, Appl	359	6	1.2	304	1	US-08-499-568-2	Sequence 2, Appl
287	6	1.2	128	1	US-08-426-677-14	Sequence 14, Appl	360	6	1.2	304	1	US-08-793-958-2	Sequence 2, Appl
288	6	1.2	133	1	US-08-268-348A-8	Sequence 8, Appl	361	6	1.2	311	3	US-08-946-214-4	Sequence 4, Appl
289	6	1.2	133	1	US-08-268-348A-10	Sequence 10, Appl	362	6	1.2	312	1	US-08-414-525-2	Sequence 2, Appl
290	6	1.2	172	2	US-08-557-309B-39	Sequence 39, Appl	363	6	1.2	315	2	US-08-484-956-91	Sequence 91, Appl
291	6	1.2	172	3	US-08-834-306-39	Sequence 39, Appl	364	6	1.2	319	4	US-08-757-653-163	Sequence 163, App
292	6	1.2	174	3	US-08-906-136A-2	Sequence 2, Appl	365	6	1.2	320	2	US-08-823-516-61	Sequence 61, Appl
293	6	1.2	174	3	US-08-906-136A-4	Sequence 4, Appl	366	6	1.2	323	3	US-08-816-977-21	Sequence 21, Appl
294	6	1.2	179	4	US-08-809-267-7	Sequence 7, Appl	367	6	1.2	326	3	US-08-816-977-25	Sequence 25, Appl
295	6	1.2	179	4	PCT-US95-13662A-7	Sequence 7, Appl	368	6	1.2	338	2	US-08-447-430A-43	Sequence 43, Appl
296	6	1.2	180	3	US-08-772-440-31	Sequence 31, Appl	369	6	1.2	343	2	US-09-123-851-1	Sequence 1, Appl
297	6	1.2	183	5	5168049-3	Patent No. 5168049	370	6	1.2	343	2	US-08-728-820-1	Sequence 1, Appl
298	6	1.2	184	3	US-09-023-082A-30	Sequence 30, Appl	371	6	1.2	344	2	US-07-941-523-24	Sequence 24, Appl
299	6	1.2	197	2	US-08-879-561-5	Sequence 5, Appl	372	6	1.2	351	2	US-08-701-191A-3	Sequence 3, Appl
300	6	1.2	198	1	US-08-642-255-131	Sequence 131, App	373	6	1.2	357	1	US-08-078-683A-8	Sequence 8, Appl
301	6	1.2	198	1	US-08-397-633A-41	Sequence 41, Appl	374	6	1.2	360	1	US-08-674-412-3	Sequence 3, Appl
302	6	1.2	198	1	US-08-397-633A-52	Sequence 52, Appl	375	6	1.2	360	1	US-08-469-421-14	Sequence 14, Appl
303	6	1.2	214	2	US-08-846-790A-1	Sequence 1, Appl	376	6	1.2	360	1	US-08-250-975-14	Sequence 14, Appl
304	6	1.2	214	3	US-08-935-333-1	Sequence 1, Appl	377	6	1.2	360	2	US-08-920-296-3	Sequence 3, Appl
305	6	1.2	218	1	US-08-644-664B-14	Sequence 14, Appl	378	6	1.2	360	2	US-08-605-002A-14	Sequence 14, Appl
306	6	1.2	218	2	US-08-761-277-14	Sequence 14, Appl	379	6	1.2	360	2	US-08-950-449A-14	Sequence 14, Appl
307	6	1.2	219	1	US-08-621-081A-16	Sequence 16, Appl	380	6	1.2	360	4	US-08-554-385-20	Sequence 20, Appl
308	6	1.2	223	1	US-08-430-633-1	Sequence 1, Appl	381	6	1.2	376	3	PCT-US94-10529-14	Sequence 14, Appl
309	6	1.2	223	2	US-08-936-854-1	Sequence 1, Appl	382	6	1.2	377	2	US-08-729-812-8	Sequence 8, Appl
310	6	1.2	224	1	US-08-360-673-7	Sequence 7, Appl	383	6	1.2	378	2	US-08-929-822B-2	Sequence 2, Appl
311	6	1.2	227	5	US-08-414-625-4	Sequence 4, Appl	384	6	1.2	378	2	US-08-707-237A-104	Sequence 104, App
312	6	1.2	227	5	5428012-4	Patent No. 5428012	385	6	1.2	378	4	PCT-US96-06229-26	Sequence 26, Appl
313	6	1.2	227	5	5428012-4	Patent No. 5428012	386	6	1.2	380	1	US-08-585-758A-4	Sequence 4, Appl
314	6	1.2	227	5	5451506-3	Patent No. 5451506	387	6	1.2	380	1	US-08-977-818-4	Sequence 4, Appl
315	6	1.2	228	5	5428012-2	Patent No. 5428012	388	6	1.2				
316	6	1.2	228	5	5451506-2	Patent No. 5451506	389	6	1.2				
317	6	1.2	233	1	US-08-446-918A-4	Sequence 4, Appl	390	6	1.2				
318	6	1.2	233	2	US-08-480-753-2	Sequence 2, Appl	391	6	1.2				
319	6	1.2	233	2	US-08-580-806-4	Sequence 4, Appl	392	6	1.2				
320	6	1.2	233	3	US-09-041-889-9	Sequence 9, Appl	393	6	1.2				

394	6	1.2	380	2	US-08-670-27AB-4	Sequence 4, Appl1	467	6	1.2	593	1	US-08-374-155A-14	Sequence 14, Appl1
395	6	1.2	381	1	US-08-585-738A-2	Sequence 2, Appl1	468	6	1.2	593	2	US-08-785-36-14	Sequence 14, Appl1
396	6	1.2	381	2	US-08-977-81B-2	Sequence 2, Appl1	469	6	1.2	601	2	US-08-606-288-7	Sequence 7, Appl1
397	6	1.2	381	2	US-08-670-27AB-2	Sequence 2, Appl1	470	6	1.2	601	2	US-08-606-288-10	Sequence 10, Appl1
398	6	1.2	381	2	US-08-786-999-3	Sequence 3, Appl1	471	6	1.2	612	2	US-08-753-307B-11	Sequence 11, Appl1
399	6	1.2	386	3	US-08-895-707-7	Sequence 7, Appl1	472	6	1.2	616	3	US-08-895-707-2	Sequence 2, Appl1
400	6	1.2	390	2	US-08-786-999-1	Sequence 1, Appl1	473	6	1.2	628	3	US-08-394-177-6	Sequence 6, Appl1
401	6	1.2	391	1	US-07-995-657-6	Sequence 6, Appl1	474	6	1.2	630	2	US-08-596-319-2	Sequence 2, Appl1
402	6	1.2	391	1	US-08-474-587-6	Sequence 6, Appl1	475	6	1.2	631	1	US-08-605-541B-12	Sequence 12, Appl1
403	6	1.2	391	2	US-08-244-951A-10	Sequence 10, Appl1	476	6	1.2	662	2	US-08-770-761A-10	Sequence 5, Appl1
404	6	1.2	391	2	US-08-389-011-23	Sequence 23, Appl1	477	6	1.2	677	3	US-08-480-640A-115	Sequence 115, App
405	6	1.2	391	3	US-08-403-917A-23	Sequence 23, Appl1	478	6	1.2	677	3	US-08-480-640A-193	Sequence 193, App
406	6	1.2	391	4	PCR-US95-15696-2	Sequence 2, Appl1	479	6	1.2	680	1	US-08-542-363-4	Sequence 4, Appl1
407	6	1.2	393	4	US-08-377-534-2	Sequence 2, Appl1	480	6	1.2	682	1	US-08-642-255-126	Sequence 126, App
408	6	1.2	401	5	5252556-1	Patent No. 5252556	481	6	1.2	682	3	US-08-397-633A-36	Sequence 36, Appl1
409	6	1.2	403	5	US-09-061-337-10	Sequence 10, Appl1	482	6	1.2	701	2	US-08-533-669A-16	Sequence 16, Appl1
410	6	1.2	403	3	US-09-122-129-10	Sequence 10, Appl1	483	6	1.2	705	2	US-08-770-761A-7	Sequence 7, Appl1
411	6	1.2	406	2	US-09-340-991-10	Sequence 10, Appl1	484	6	1.2	729	1	US-07-640-039-3	Sequence 3, Appl1
412	6	1.2	406	2	US-08-222-719-1	Sequence 1, Appl1	485	6	1.2	731	1	US-07-921-807B-5	Sequence 5, Appl1
413	6	1.2	406	2	US-08-470-925-1	Sequence 1, Appl1	486	6	1.2	731	1	US-08-441-944A-5	Sequence 5, Appl1
414	6	1.2	406	2	US-08-471-613-1	Sequence 1, Appl1	487	6	1.2	733	1	US-07-640-029-4	Sequence 4, Appl1
415	6	1.2	406	2	PCR-US93-10443-1	Sequence 1, Appl1	488	6	1.2	733	1	US-07-921-807B-6	Sequence 6, Appl1
416	6	1.2	407	2	US-08-776-585-3	Sequence 3, Appl1	489	6	1.2	733	1	US-08-441-944A-6	Sequence 6, Appl1
417	6	1.2	423	2	US-08-910-798-2	Sequence 2, Appl1	490	6	1.2	733	1	US-08-188-582-16	Sequence 16, Appl1
418	6	1.2	423	2	US-08-313-553-13	Sequence 13, Appl1	491	6	1.2	737	1	US-08-646-715-16	Sequence 16, Appl1
419	6	1.2	423	3	US-08-767-993-13	Sequence 13, Appl1	492	6	1.2	737	1	US-08-646-715-16	Sequence 16, Appl1
420	6	1.2	427	2	US-08-846-021A-8	Sequence 8, Appl1	493	6	1.2	741	2	US-08-849-480A-9	Sequence 9, Appl1
421	6	1.2	439	1	US-08-507-431-4	Sequence 4, Appl1	494	6	1.2	761	2	US-08-707-237A-84	Sequence 84, Appl1
422	6	1.2	439	2	US-08-700-546-2	Sequence 2, Appl1	495	6	1.2	762	1	US-08-642-255-114	Sequence 114, App
423	6	1.2	439	2	US-08-902-655A-4	Sequence 4, Appl1	496	6	1.2	762	1	US-08-642-255-120	Sequence 120, App
424	6	1.2	439	3	US-09-116-622-4	Sequence 4, Appl1	497	6	1.2	762	1	US-08-397-633A-26	Sequence 26, Appl1
425	6	1.2	450	2	US-08-665-037-2	Sequence 2, Appl1	498	6	1.2	762	1	US-08-397-633A-31	Sequence 31, Appl1
426	6	1.2	450	2	US-08-666-067-2	Sequence 2, Appl1	499	6	1.2	764	1	US-08-424-567-2	Sequence 2, Appl1
427	6	1.2	450	2	US-08-732-870-2	Sequence 2, Appl1	500	6	1.2	764	2	US-08-711-928-2	Sequence 2, Appl1
428	6	1.2	458	5	5217891-15	Patent No. 5217891	501	6	1.2	768	2	US-08-560-398-2	Sequence 2, Appl1
429	6	1.2	474	2	US-09-096-982-8	Sequence 8, Appl1	502	6	1.2	770	1	US-08-395-706-12	Sequence 12, Appl1
430	6	1.2	474	2	US-08-653-650A-8	Sequence 8, Appl1	503	6	1.2	770	1	US-08-416-581B-1	Sequence 1, Appl1
431	6	1.2	479	1	US-08-397-633A-78	Sequence 78, Appl1	504	6	1.2	770	1	US-08-416-581B-5	Sequence 5, Appl1
432	6	1.2	481	2	US-08-477-451-19	Sequence 19, Appl1	505	6	1.2	770	1	US-08-416-581B-9	Sequence 9, Appl1
433	6	1.2	486	1	US-08-397-633A-77	Sequence 77, Appl1	506	6	1.2	770	2	US-08-853-091-12	Sequence 12, Appl1
434	6	1.2	494	4	US-08-464-340A-4	Sequence 4, Appl1	507	6	1.2	770	3	US-08-820-754-12	Sequence 12, Appl1
435	6	1.2	494	4	PCR-US94-08449A-4	Sequence 4, Appl1	508	6	1.2	770	3	US-08-956-652-12	Sequence 12, Appl1
436	6	1.2	498	2	US-07-816-098A-4	Sequence 4, Appl1	509	6	1.2	770	4	PCR-US95-17025-12	Sequence 12, Appl1
437	6	1.2	500	2	US-08-578-709-15	Sequence 15, Appl1	510	6	1.2	771	1	US-08-276-099A-14	Sequence 14, Appl1
438	6	1.2	504	2	US-08-896-005-3	Sequence 3, Appl1	511	6	1.2	771	1	US-08-781-890-14	Sequence 14, Appl1
439	6	1.2	511	3	US-09-105-039A-2	Sequence 2, Appl1	512	6	1.2	777	2	US-08-477-396A-4	Sequence 4, Appl1
440	6	1.2	514	1	US-08-865-311-2	Sequence 8, Appl1	513	6	1.2	777	1	US-08-426-637-4	Sequence 4, Appl1
441	6	1.2	519	1	US-08-589-446-8	Sequence 8, Appl1	514	6	1.2	779	1	US-08-426-637-24	Sequence 24, Appl1
442	6	1.2	519	2	US-08-444-882-8	Sequence 8, Appl1	515	6	1.2	780	1	US-08-375-709-13	Sequence 13, Appl1
443	6	1.2	519	2	US-08-389-459A-8	Sequence 8, Appl1	516	6	1.2	780	1	US-08-752-929-13	Sequence 13, Appl1
444	6	1.2	519	3	US-08-987-867A-8	Sequence 8, Appl1	517	6	1.2	788	1	US-08-194-338-12	Sequence 12, Appl1
445	6	1.2	524	1	US-08-447-500-24	Sequence 24, Appl1	518	6	1.2	788	1	US-08-572-225-1	Sequence 1, Appl1
446	6	1.2	524	1	US-08-454-097-24	Sequence 24, Appl1	519	6	1.2	811	1	US-08-426-637-2	Sequence 2, Appl1
447	6	1.2	524	1	US-08-453-866-24	Sequence 24, Appl1	520	6	1.2	811	1	US-08-426-637-22	Sequence 22, Appl1
448	6	1.2	525	3	US-08-185-359-24	Sequence 24, Appl1	521	6	1.2	816	1	US-07-640-029-1	Sequence 1, Appl1
449	6	1.2	525	1	US-08-375-709-19	Sequence 19, Appl1	522	6	1.2	817	1	US-07-640-029-2	Sequence 2, Appl1
450	6	1.2	525	1	US-08-752-929-19	Sequence 19, Appl1	523	6	1.2	820	1	US-07-921-807B-3	Sequence 3, Appl1
451	6	1.2	526	2	US-08-852-401-3	Sequence 3, Appl1	524	6	1.2	820	1	US-08-291-896-2	Sequence 2, Appl1
452	6	1.2	527	1	US-08-160-670A-8	Sequence 8, Appl1	525	6	1.2	820	1	US-08-441-944A-3	Sequence 3, Appl1
453	6	1.2	530	3	US-09-105-039A-4	Sequence 4, Appl1	526	6	1.2	820	1	US-08-166-717D-6	Sequence 6, Appl1
454	6	1.2	530	3	US-08-737-336-6	Sequence 6, Appl1	527	6	1.2	820	2	US-08-485-278-2	Sequence 2, Appl1
455	6	1.2	537	3	US-08-920-610-2	Sequence 2, Appl1	528	6	1.2	822	1	US-07-997-807B-4	Sequence 4, Appl1
456	6	1.2	539	3	US-08-906-360-1	Sequence 1, Appl1	529	6	1.2	822	1	US-07-921-807B-4	Sequence 4, Appl1
457	6	1.2	547	1	US-08-785-048-3	Sequence 3, Appl1	530	6	1.2	822	1	US-08-459-286-2	Sequence 2, Appl1
458	6	1.2	547	1	US-08-996-799-3	Sequence 3, Appl1	531	6	1.2	822	1	US-08-441-944A-4	Sequence 4, Appl1
459	6	1.2	550	4	PCR-US93-03027-1	Sequence 1, Appl1	532	6	1.2	822	2	US-08-451-822A-12	Sequence 12, Appl1
460	6	1.2	553	1	US-08-785-048-2	Sequence 2, Appl1	533	6	1.2	829	1	US-08-642-255-132	Sequence 132, App
461	6	1.2	553	2	US-08-661-052-16	Sequence 16, Appl1	534	6	1.2	829	3	US-08-397-633A-53	Sequence 53, Appl1
462	6	1.2	553	2	US-08-996-799-2	Sequence 2, Appl1	535	6	1.2	832	1	US-08-630-820-7	Sequence 7, Appl1
463	6	1.2	559	1	US-08-320-559-31	Sequence 31, Appl1	536	6	1.2	836	1	US-08-426-637-6	Sequence 6, Appl1
464	6	1.2	559	1	US-08-545-860D-31	Sequence 31, Appl1	537	6	1.2	837	1	US-08-426-637-23	Sequence 23, Appl1
465	6	1.2	559	4	PCR-US94-04496-31	Sequence 31, Appl1	538	6	1.2	836	2	US-07-916-098A-2	Sequence 2, Appl1
466	6	1.2	561	1	US-08-360-673-2	Sequence 2, Appl1	539	6	1.2	870	2	US-08-732-192A-2	Sequence 2, Appl1

540	6	1.2	872	2	US-08-844-057-2	Sequence 2, Appl1	613	6	1.2	2105	2	US-08-808-793-3	Sequence 3, Appl1
541	6	1.2	877	1	US-08-397-633a-54	Sequence 54, Appl1	614	6	1.2	2105	3	US-08-772-512A-3	Sequence 3, Appl1
542	6	1.2	878	1	US-08-37-919-2	Sequence 2, Appl1	615	6	1.2	2206	1	US-07-852-260-2	Sequence 2, Appl1
543	6	1.2	878	4	PCT-US95-0518-2	Sequence 2, Appl1	616	6	1.2	2206	2	US-08-461-503-2	Sequence 2, Appl1
544	6	1.2	884	2	US-08-397-633a-68	Sequence 68, Appl1	617	6	1.2	2842	1	US-07-741-940-7	Sequence 7, Appl1
545	6	1.2	884	2	US-08-474-067-8	Sequence 8, Appl1	618	6	1.2	2842	1	US-08-289-548A-7	Sequence 7, Appl1
546	6	1.2	884	2	US-08-435-641-15	Sequence 15, Appl1	619	6	1.2	2842	1	US-08-452-654-7	Sequence 7, Appl1
547	6	1.2	884	2	US-08-707-337a-96	Sequence 96, Appl1	620	6	1.2	2843	1	US-07-741-940-2	Sequence 2, Appl1
548	6	1.2	884	2	US-08-474-068a-8	Sequence 8, Appl1	621	6	1.2	2843	1	US-08-289-548A-2	Sequence 2, Appl1
549	6	1.2	884	2	US-08-472-481-7	Sequence 7, Appl1	622	6	1.2	2843	1	US-08-452-654-2	Sequence 2, Appl1
550	6	1.2	884	3	US-08-642-246-15	Sequence 15, Appl1	623	6	1.2	2843	1	US-08-452-653B-2	Sequence 2, Appl1
551	6	1.2	884	4	PCT-US96-0622a-15	Sequence 15, Appl1	624	6	1.2	2843	1	US-08-452-653B-7	Sequence 7, Appl1
552	6	1.2	933	2	US-08-313-200-1	Sequence 1, Appl1	625	6	1.2	2843	2	US-08-370-235A-2	Sequence 2, Appl1
553	6	1.2	933	4	PCT-US93-03837-1	Sequence 1, Appl1	626	6	1.2	2973	2	US-08-821-355A-7	Sequence 7, Appl1
554	6	1.2	936	2	US-08-707-237a-108	Sequence 108, App	627	6	1.2	2973	2	US-09-003-687A-7	Sequence 7, Appl1
555	6	1.2	936	3	US-08-642-246-30	Sequence 30, Appl1	628	6	1.2	3691	2	US-07-642-734C-4	Sequence 2, Appl1
556	6	1.2	936	4	PCT-US96-0622a-30	Sequence 30, Appl1	629	6	1.2	3691	2	US-07-642-734C-4	Sequence 2, Appl1
557	6	1.2	966	3	US-08-642-246-34	Sequence 34, Appl1	630	6	1.2	3724	2	US-08-804-227C-10	Sequence 10, Appl1
558	6	1.2	966	4	PCT-US96-0622a-34	Sequence 34, Appl1	631	6	1.2	3724	2	US-08-804-198-4	Sequence 4, Appl1
559	6	1.2	970	2	US-08-449-645a-11	Sequence 11, Appl1	632	6	1.2	15281	2	US-08-471-119A-2	Sequence 2, Appl1
560	6	1.2	970	2	US-08-702-367a-11	Sequence 11, Appl1	633	6	1.0	5	1	US-07-816-678A-7	Sequence 7, Appl1
561	6	1.2	980	4	PCT-US95-04681-11	Sequence 11, Appl1	634	6	1.0	5	1	US-07-826-928A-7	Sequence 7, Appl1
562	6	1.2	980	1	US-08-091-569-2	Sequence 2, Appl1	635	6	1.0	5	1	US-08-197-496A-21	Sequence 21, Appl1
563	6	1.2	980	1	US-08-220-151-5	Sequence 5, Appl1	636	6	1.0	5	1	US-08-131-365B-49	Sequence 49, Appl1
564	6	1.2	980	1	US-08-403-676-2	Sequence 2, Appl1	637	6	1.0	5	1	US-08-275-370-26	Sequence 26, Appl1
565	6	1.2	980	1	US-08-413-118-5	Sequence 5, Appl1	638	6	1.0	5	1	US-08-367-968-26	Sequence 26, Appl1
566	6	1.2	980	2	US-08-822-238-2	Sequence 2, Appl1	639	6	1.0	5	1	US-08-294-434-2	Sequence 2, Appl1
567	6	1.2	980	3	US-08-473-446-5	Sequence 5, Appl1	640	6	1.0	5	1	US-08-457-166-2	Sequence 2, Appl1
568	6	1.2	990	2	US-08-392-625-20	Sequence 20, Appl1	641	6	1.0	5	1	US-08-200-900A-34	Sequence 34, Appl1
569	6	1.2	990	2	US-08-466-961a-20	Sequence 20, Appl1	642	6	1.0	5	1	US-08-480-604A-32	Sequence 32, Appl1
570	6	1.2	990	2	US-08-645-193a-15	Sequence 15, Appl1	643	6	1.0	5	2	US-08-665-484-26	Sequence 26, Appl1
571	6	1.2	995	1	US-08-162-809-18	Sequence 18, Appl1	644	6	1.0	5	2	US-09-016-366A-29	Sequence 29, Appl1
572	6	1.2	995	2	US-08-673-789-5	Sequence 5, Appl1	645	6	1.0	5	2	US-07-963-538B-11	Sequence 11, Appl1
573	6	1.2	995	2	US-08-717-515-8	Sequence 14, Appl1	646	6	1.0	5	2	US-08-623-833B-55	Sequence 55, Appl1
574	6	1.2	995	4	PCT-US95-04910-14	Sequence 103, App	647	6	1.0	5	2	US-08-668-122-45	Sequence 45, Appl1
575	6	1.2	1002	2	US-08-707-237a-103	Sequence 25, Appl1	648	6	1.0	5	2	US-08-405-496A-30	Sequence 30, Appl1
576	6	1.2	1002	3	US-08-642-246-25	Sequence 25, Appl1	649	6	1.0	5	2	US-08-595-043A-13	Sequence 13, Appl1
577	6	1.2	1002	4	PCT-US96-0622a-25	Sequence 12, Appl1	650	6	1.0	5	2	US-09-016-366A-29	Sequence 29, Appl1
578	6	1.2	1011	1	US-08-162-809-12	Sequence 2, Appl1	651	6	1.0	5	3	US-08-595-868C-16	Sequence 16, Appl1
579	6	1.2	1046	1	US-08-386-727-2	Sequence 2, Appl1	652	6	1.0	5	3	US-08-481-435-41	Sequence 41, Appl1
580	6	1.2	1084	1	US-08-600-452a-2	Sequence 2, Appl1	653	6	1.0	5	3	US-08-483-748A-33	Sequence 33, Appl1
581	6	1.2	1194	1	US-08-717-515-6	Sequence 6, Appl1	654	6	1.0	5	4	US-08-997-918-56	Sequence 56, Appl1
582	6	1.2	1220	2	US-08-680-326-35	Sequence 35, Appl1	655	6	1.0	5	4	PCT-US92-11270-7	Sequence 7, Appl1
583	6	1.2	1220	2	US-08-80-326-38	Sequence 38, Appl1	656	6	1.0	5	4	PCT-US93-06591-2	Sequence 2, Appl1
584	6	1.2	1253	1	US-08-252-966B-12	Sequence 12, Appl1	657	6	1.0	5	4	US-08-923-15800-2	Sequence 21, Appl1
585	6	1.2	1261	1	US-08-352-966B-18	Sequence 18, Appl1	658	6	1.0	6	1	US-08-243-082-21	Sequence 2, Appl1
586	6	1.2	1276	1	US-08-717-515-8	Sequence 8, Appl1	659	6	1.0	6	1	US-08-343-943-2	Sequence 2, Appl1
587	6	1.2	1394	4	PCT-US95-10661A-2	Sequence 2, Appl1	660	6	1.0	6	1	US-08-115-519-6	Sequence 6, Appl1
588	6	1.2	1398	1	US-08-750-532-9	Sequence 9, Appl1	661	6	1.0	6	2	US-08-200-900A-35	Sequence 35, Appl1
589	6	1.2	1452	2	US-08-652-971-4	Sequence 4, Appl1	662	6	1.0	6	2	US-07-885-089B-19	Sequence 19, Appl1
590	6	1.2	1452	2	US-08-991-258A-4	Sequence 4, Appl1	663	6	1.0	6	2	US-08-531-525-1	Sequence 1, Appl1
591	6	1.2	1452	2	US-08-769-399-4	Sequence 4, Appl1	664	6	1.0	6	2	US-07-963-538B-6	Sequence 6, Appl1
592	6	1.2	1477	1	US-08-038-682-4	Sequence 4, Appl1	665	6	1.0	6	2	US-08-692-825-39	Sequence 39, Appl1
593	6	1.2	1477	1	US-08-302-832-4	Sequence 4, Appl1	666	6	1.0	6	2	US-08-718-270A-1	Sequence 1, Appl1
594	6	1.2	1477	2	US-08-530-198-4	Sequence 4, Appl1	667	6	1.0	6	2	US-08-845-742-8	Sequence 8, Appl1
595	6	1.2	1477	2	US-08-469-880-4	Sequence 4, Appl1	668	6	1.0	7	1	US-08-276-855-14	Sequence 14, Appl1
596	6	1.2	1477	2	US-08-728-470-4	Sequence 4, Appl1	669	6	1.0	7	1	US-08-899-575-14	Sequence 14, Appl1
597	6	1.2	1477	2	US-08-617-697-4	Sequence 4, Appl1	670	6	1.0	7	1	US-08-115-519-7	Sequence 7, Appl1
598	6	1.2	1541	4	PCT-US95-10661A-3	Sequence 3, Appl1	671	6	1.0	7	1	US-08-704-176-43	Sequence 43, Appl1
599	6	1.2	1566	2	US-08-687-956A-23	Sequence 23, Appl1	672	6	1.0	7	2	US-08-895-648-34	Sequence 34, Appl1
600	6	1.2	1620	1	US-08-542-363-2	Sequence 2, Appl1	673	6	1.0	7	3	US-08-951-648-34	Sequence 34, Appl1
601	6	1.2	1621	1	US-08-242-677-2	Sequence 2, Appl1	674	6	1.0	7	3	US-08-481-435-28	Sequence 28, Appl1
602	6	1.2	1627	1	US-07-665-792E-9	Sequence 9, Appl1	675	6	1.0	7	3	PCT-US94-02631-43	Sequence 43, Appl1
603	6	1.2	1648	2	US-08-662-227-35	Sequence 35, Appl1	676	6	1.0	7	4	PCT-US95-08743-14	Sequence 14, Appl1
604	6	1.2	1664	2	US-08-642-845-2	Sequence 2, Appl1	677	6	1.0	8	1	US-08-458-458B-7	Sequence 7, Appl1
605	6	1.2	1820	2	US-07-998-289E-8	Sequence 8, Appl1	678	6	1.0	8	1	US-08-323-474-4	Sequence 4, Appl1
606	6	1.2	1841	2	US-08-804-227C-6	Sequence 6, Appl1	679	6	1.0	8	1	US-08-225-989-15	Sequence 15, Appl1
607	6	1.2	1958	1	US-07-945-283-2	Sequence 23, Appl1	680	6	1.0	8	1	US-08-086-335C-12	Sequence 12, Appl1
608	6	1.2	2089	1	US-08-418-893D-23	Sequence 23, Appl1	681	6	1.0	8	1	US-07-982-174-1	Sequence 1, Appl1
609	6	1.2	2100	3	US-08-418-893D-24	Sequence 23, Appl1	682	6	1.0	8	1	US-08-263-086-5	Sequence 5, Appl1
610	6	1.2	2100	3	US-08-808-793-23	Sequence 23, Appl1	683	6	1.0	8	1	US-08-318-193-30	Sequence 30, Appl1
611	6	1.2	2104	2	US-08-772-512A-19	Sequence 4, Appl1	684	6	1.0	8	1	US-08-240-514-34	Sequence 34, Appl1
612	6	1.2	2104	3	US-08-808-793-4	Sequence 4, Appl1	685	6	1.0	8	1	US-08-236-918A-16	Sequence 16, Appl1

666	5	1.0	8	1	US-08-570-923-15	Sequence 15, Appl	759	1.0	8	4	PCT-US95-00362-3	Sequence 3, Appl
667	5	1.0	8	1	US-08-446-908-6	Sequence 6, Appl	760	1.0	8	4	PCT-US95-06530-7	Sequence 7, Appl
668	5	1.0	8	1	US-08-297-706-1	Sequence 1, Appl	761	1.0	8	4	PCT-US95-08534-3	Sequence 3, Appl
669	5	1.0	8	1	US-08-634-060-25	Sequence 25, Appl	762	1.0	8	4	PCT-US95-10973A-98	Sequence 98, Appl
670	5	1.0	8	1	US-08-231-205A-6	Sequence 6, Appl	763	1.0	8	4	PCT-US95-11960-2	Sequence 2, Appl
671	5	1.0	8	1	US-08-446-922-1	Sequence 1, Appl	764	1.0	8	4	PCT-US95-115781-6	Sequence 6, Appl
672	5	1.0	8	1	US-08-430-633-3	Sequence 3, Appl	765	1.0	8	4	PCT-US95-11800-19	Sequence 19, Appl
673	5	1.0	8	1	US-08-580-014-15	Sequence 15, Appl	766	1.0	8	4	PCT-US95-10895-7	Sequence 7, Appl
674	5	1.0	8	1	US-08-461-597-8	Sequence 8, Appl	767	1.0	8	5	5194375-7	Patent No. 5194375
675	5	1.0	8	1	US-08-670-354-7	Sequence 7, Appl	768	1.0	8	5	5198342-3	Patent No. 5198342
676	5	1.0	8	1	US-08-442-043A-15	Sequence 15, Appl	769	1.0	8	5	5395760-6	Patent No. 5395760
677	5	1.0	8	1	US-08-604-333-5	Sequence 5, Appl	770	1.0	9	1	US-08-787-547-98	Sequence 98, Appl
678	5	1.0	8	1	US-08-308-881-7	Sequence 7, Appl	771	1.0	9	2	US-08-340-283-134	Sequence 134, Appl
679	5	1.0	8	1	US-08-660-626-1	Sequence 1, Appl	772	1.0	9	2	US-08-605-002A-17	Sequence 17, Appl
700	5	1.0	8	1	US-08-526-014-1	Sequence 1, Appl	773	1.0	9	2	US-08-595-043A-41	Sequence 41, Appl
701	5	1.0	8	2	US-08-612-302A-34	Sequence 34, Appl	774	1.0	9	2	US-08-889-931A-3	Sequence 3, Appl
702	5	1.0	8	2	US-08-588-528-1	Sequence 1, Appl	775	1.0	9	2	US-08-960-449A-19	Sequence 19, Appl
703	5	1.0	8	2	US-08-684-687-6	Sequence 6, Appl	776	1.0	9	2	US-08-986-234-46	Sequence 46, Appl
704	5	1.0	8	2	US-08-701-124-2	Sequence 2, Appl	777	1.0	9	3	US-08-968-747-4	Sequence 4, Appl
705	5	1.0	8	2	US-08-649-341A-9	Sequence 9, Appl	778	1.0	9	3	US-08-997-918-53	Sequence 53, Appl
706	5	1.0	8	2	US-08-494-440B-13	Sequence 13, Appl	779	1.0	10	1	US-07-721-761A-16	Sequence 16, Appl
707	5	1.0	8	2	US-08-729-345-4	Sequence 4, Appl	780	1.0	10	1	US-07-978-687-16	Sequence 16, Appl
708	5	1.0	8	2	US-08-533-901B-15	Sequence 15, Appl	781	1.0	10	1	US-08-435-801-26	Sequence 26, Appl
709	5	1.0	8	2	US-08-670-175-1	Sequence 1, Appl	782	1.0	10	1	US-08-568-072-1	Sequence 1, Appl
710	5	1.0	8	2	US-08-777-405A-15	Sequence 15, Appl	783	1.0	10	1	US-08-603-753D-13	Sequence 13, Appl
711	5	1.0	8	2	US-08-922-267A-63	Sequence 63, Appl	784	1.0	10	2	US-08-603-753D-14	Sequence 14, Appl
712	5	1.0	8	2	US-08-535-298-10	Sequence 10, Appl	785	1.0	10	2	US-08-603-753D-15	Sequence 15, Appl
713	5	1.0	8	2	US-08-620-694A-3	Sequence 3, Appl	786	1.0	10	2	US-08-603-753D-16	Sequence 16, Appl
714	5	1.0	8	2	US-08-720-258-7	Sequence 7, Appl	787	1.0	10	2	US-08-603-753D-17	Sequence 17, Appl
715	5	1.0	8	2	US-08-977-871A-15	Sequence 15, Appl	788	1.0	10	2	US-08-713-928B-14	Sequence 14, Appl
716	5	1.0	8	2	US-08-839-032A-19	Sequence 19, Appl	789	1.0	10	2	US-08-595-668C-49	Sequence 49, Appl
717	5	1.0	8	2	US-08-058-263-7	Sequence 7, Appl	790	1.0	10	3	US-08-467-580-39	Sequence 39, Appl
718	5	1.0	8	2	US-08-870-518-27	Sequence 27, Appl	791	1.0	10	3	US-08-159-339A-295	Sequence 295, Appl
719	5	1.0	8	2	US-08-936-854-3	Sequence 3, Appl	792	1.0	10	4	PCT-US91-339A-376	Sequence 376, Appl
720	5	1.0	8	2	US-09-059-099-7	Sequence 7, Appl	793	1.0	10	4	PCT-US91-05801-16	Sequence 16, Appl
721	5	1.0	8	2	US-08-713-928B-10	Sequence 10, Appl	794	1.0	10	4	PCT-US93-04717-26	Sequence 26, Appl
722	5	1.0	8	2	US-08-595-043A-1	Sequence 1, Appl	795	1.0	10	4	PCT-US95-08516-39	Sequence 39, Appl
723	5	1.0	8	2	US-08-690-011A-50	Sequence 50, Appl	796	1.0	10	4	PCT-US95-12960-1	Sequence 1, Appl
724	5	1.0	8	2	US-08-839-031A-15	Sequence 15, Appl	797	1.0	11	1	US-08-061-350-2	Sequence 2, Appl
725	5	1.0	8	2	US-08-762-106-3	Sequence 3, Appl	798	1.0	11	1	US-08-200-000A-37	Sequence 37, Appl
726	5	1.0	8	2	US-09-016-366A-30	Sequence 30, Appl	799	1.0	11	1	US-08-338-634-34	Sequence 34, Appl
727	5	1.0	8	2	US-08-700-846-8	Sequence 8, Appl	800	1.0	11	1	US-08-471-780C-120	Sequence 120, Appl
728	5	1.0	8	2	US-08-871-161-6	Sequence 6, Appl	801	1.0	11	1	US-08-471-780C-124	Sequence 124, Appl
729	5	1.0	8	2	US-08-832-935-1	Sequence 1, Appl	802	1.0	11	1	US-08-467-282B-120	Sequence 120, Appl
730	5	1.0	8	2	US-08-482-728A-2	Sequence 2, Appl	803	1.0	11	1	US-08-467-282B-124	Sequence 124, Appl
731	5	1.0	8	2	US-08-826-989-5	Sequence 5, Appl	804	1.0	11	2	US-08-471-282A-120	Sequence 120, Appl
732	5	1.0	8	2	US-09-225-951-15	Sequence 15, Appl	805	1.0	11	2	US-08-471-282A-124	Sequence 124, Appl
733	5	1.0	8	3	US-09-058-264-7	Sequence 7, Appl	806	1.0	11	2	US-08-318-157B-40	Sequence 40, Appl
734	5	1.0	8	3	US-08-991-426-8	Sequence 8, Appl	807	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
735	5	1.0	8	3	US-08-996-139-7	Sequence 7, Appl	808	1.0	11	2	US-08-466-710C-124	Sequence 124, Appl
736	5	1.0	8	3	US-08-792-832A-7	Sequence 7, Appl	809	1.0	11	2	US-08-560-558E-12	Sequence 8, Appl
737	5	1.0	8	3	US-09-130-663-6	Sequence 6, Appl	810	1.0	11	2	US-08-755-288-18	Sequence 12, Appl
738	5	1.0	8	3	US-09-081-180-16	Sequence 16, Appl	811	1.0	11	2	US-08-974-655-12	Sequence 12, Appl
739	5	1.0	8	3	US-09-040-786-16	Sequence 16, Appl	812	1.0	11	3	US-08-468-739C-120	Sequence 120, Appl
740	5	1.0	8	3	US-08-819-177-14	Sequence 14, Appl	813	1.0	11	3	US-08-468-739C-124	Sequence 124, Appl
741	5	1.0	8	3	US-08-897-924A-20	Sequence 20, Appl	814	1.0	11	3	US-09-075-257A-4	Sequence 4, Appl
742	5	1.0	8	3	US-08-307-896-9	Sequence 9, Appl	815	1.0	11	3	US-09-075-257A-4	Sequence 4, Appl
743	5	1.0	8	3	US-08-718-904-32	Sequence 32, Appl	816	1.0	11	3	US-08-913-264-8	Sequence 8, Appl
744	5	1.0	8	3	US-08-785-247-9	Sequence 9, Appl	817	1.0	11	3	US-08-350-884-53	Sequence 53, Appl
745	5	1.0	8	3	US-08-938-830-8	Sequence 8, Appl	818	1.0	12	1	US-08-709-173-53	Sequence 53, Appl
746	5	1.0	8	3	US-08-828-741B-9	Sequence 9, Appl	819	1.0	12	2	US-08-701-124-33	Sequence 23, Appl
747	5	1.0	8	3	US-08-819-177-14	Sequence 14, Appl	820	1.0	12	2	US-08-709-177-53	Sequence 53, Appl
748	5	1.0	8	3	US-09-143-470-18	Sequence 18, Appl	821	1.0	12	3	US-08-792-832A-9	Sequence 9, Appl
749	5	1.0	8	3	US-08-950-720A-8	Sequence 8, Appl	822	1.0	12	3	US-08-130-225-23	Sequence 23, Appl
750	5	1.0	8	3	US-09-130-225-2	Sequence 2, Appl	823	1.0	13	1	US-08-222-616-41	Sequence 41, Appl
751	5	1.0	8	3	US-09-022-253-3	Sequence 3, Appl	824	1.0	13	1	US-08-306-231-6	Sequence 6, Appl
752	5	1.0	8	3	US-08-883-036A-5	Sequence 5, Appl	825	1.0	13	1	US-08-446-908-7	Sequence 7, Appl
753	5	1.0	8	3	PCT-US93-05759-1	Sequence 1, Appl	826	1.0	13	1	US-08-231-205A-7	Sequence 22, Appl
754	5	1.0	8	4	PCT-US93-10034-1	Sequence 1, Appl	827	1.0	13	1	US-08-355-888A-22	Sequence 22, Appl
755	5	1.0	8	4	PCT-US93-11492-1	Sequence 1, Appl	828	1.0	13	2	US-08-693-697-12	Sequence 22, Appl
756	5	1.0	8	4	PCT-US94-08119-15	Sequence 15, Appl	829	1.0	13	2	US-08-690-011A-48	Sequence 48, Appl
757	5	1.0	8	4	PCT-US94-12913A-15	Sequence 15, Appl	830	1.0	13	2	US-08-871-161-7	Sequence 7, Appl
758	5	1.0	8	4	PCT-US94-12913A-15	Sequence 15, Appl	831	1.0	13	2	US-08-860-174A-30	Sequence 30, Appl

978 5 1.0 27 2 US-07-684-965-7 Sequence 7, Appl1
 979 5 1.0 27 2 US-07-684-965-11 Sequence 11, Appl1
 980 5 1.0 27 3 US-08-433-522A-44 Sequence 44, Appl1
 981 5 1.0 27 3 US-09-276-202-2 Sequence 2, Appl1
 982 5 1.0 27 3 US-09-276-202-8 Sequence 8, Appl1
 983 5 1.0 28 1 US-07-927-071-5 Sequence 5, Appl1
 984 5 1.0 28 2 US-08-448-603A-9 Sequence 9, Appl1
 985 5 1.0 28 3 US-08-942-423-9 Sequence 9, Appl1
 986 5 1.0 28 3 US-09-134-075-9 Sequence 9, Appl1
 987 5 1.0 29 1 US-07-681-701-9 Sequence 9, Appl1
 988 5 1.0 29 1 US-08-053-131-73 Sequence 73, Appl1
 989 5 1.0 29 1 US-08-645-641-73 Sequence 73, Appl1
 990 5 1.0 29 1 US-07-853-408B-73 Sequence 73, Appl1
 991 5 1.0 29 2 US-08-096-762-73 Sequence 73, Appl1
 992 5 1.0 29 2 US-08-308-865-73 Sequence 73, Appl1
 993 5 1.0 29 4 PCT-US93-10983-73 Sequence 2, Appl1
 994 5 1.0 30 1 US-08-273-776-2 Sequence 2, Appl1
 995 5 1.0 30 1 US-08-294-434-6 Sequence 6, Appl1
 996 5 1.0 30 1 US-08-457-166-6 Sequence 6, Appl1
 997 5 1.0 30 1 US-08-634-060-39 Sequence 39, Appl1
 998 5 1.0 30 2 US-08-757-316C-31 Sequence 31, Appl1
 999 5 1.0 30 2 US-08-143-311B-9 Sequence 9, Appl1
 1000 5 1.0 30 2 US-08-480-473B-35 Sequence 35, Appl1

ALIGNMENTS

RESULT 1
 US-08-913-477-2
 ; Sequence 2, Application US/08913477
 ; Patent No. 5985285
 ; GENERAL INFORMATION:
 ; APPLICANT: Titball, Richard W.
 ; APPLICANT: Williamson, Ethel D.
 ; APPLICANT: Leary, Sophie E.C.
 ; APPLICANT: Oyston, Petra C.F.
 ; APPLICANT: Bennett, Alice M.
 ; TITLE OF INVENTION: VACCINES FOR PLAGUE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHAYE P.C.
 ; STREET: 1100 NO. 5985285th Gleebe Rd. 8th floor
 ; CITY: Arlington
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22201-4741
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/913,477
 ; FILING DATE: 15-SEP-1997
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB96/00571
 ; FILING DATE: 13-MAR-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9505059.7
 ; FILING DATE: 13-MAR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9518946.0
 ; FILING DATE: 15-SEP-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9524825.8
 ; FILING DATE: 05-DEC-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crawford, Arthur R.
 ; REGISTRATION NUMBER: 25,327
 ; REFERENCE/DOCKET NUMBER: 124-599
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000
 TELEFAX: 703-816-4100
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 329 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-913-477-2
 Query Match 46.6%; Score 243; DB 2; Length 329;
 Best Local Similarity 100.0%; Pred. No. 8.3e-229;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 279 ILKGYHDNOLQNGIKRKEFLSPNTOWELAFYAVMHFSLTADRIDDIKYVDSM 338
 |||
 DB 87 ILKGYHDNOLQNGIKRKEFLSPNTOWELAFYAVMHFSLTADRIDDIKYVDSM 146
 QY 339 NHHGDARSKURELAEITLAEIKYVYQAEINKHLSGGTINIHDSINLMDKNLYGYTD 398
 |||
 DB 147 NHHGDARSKURELAEITLAEIKYVYQAEINKHLSGGTINIHDSINLMDKNLYGYTD 206
 QY 399 EEIFKASAEYKILEKMPQTTIYDVGSEKTIYSIKDPLGSENKRTGALGNKSYSNKDN 458
 |||
 DB 207 EEIFKASAEYKILEKMPQTTIYDVGSEKTIYSIKDPLGSENKRTGALGNKSYSNKDN 266
 QY 459 NELSHFATTCSDKSRPLNDLVSOQTOLSDITSRFSNAIENLNRFLQKDYVQORLLDPT 518
 |||
 DB 267 NELSHFATTCSDKSRPLNDLVSOQTOLSDITSRFSNAIENLNRFLQKDYVQORLLDPT 326
 QY 519 SGR 521
 |||
 DB 327 SGK 329
 RESULT 2
 US-08-913-477-4
 ; Sequence 4, Application US/08913477
 ; Patent No. 5985285
 ; GENERAL INFORMATION:
 ; APPLICANT: Titball, Richard W.
 ; APPLICANT: Williamson, Ethel D.
 ; APPLICANT: Leary, Sophie E.C.
 ; APPLICANT: Oyston, Petra C.F.
 ; APPLICANT: Bennett, Alice M.
 ; TITLE OF INVENTION: VACCINES FOR PLAGUE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHAYE P.C.
 ; STREET: 1100 NO. 5985285th Gleebe Rd. 8th floor
 ; CITY: Arlington
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22201-4741
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/913,477
 ; FILING DATE: 15-SEP-1997
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB96/00571
 ; FILING DATE: 13-MAR-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9505059.7
 ; FILING DATE: 13-MAR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9518946.0
 ; FILING DATE: 15-SEP-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-4

Query Match 46.6%; Score 243; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 8.3e-229;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHYDNLONGIRKRVKFELESPTQWELRAFAVAMHFSLTADRIDDILKVIYDSM 338
DB 87 ILKGGHYDNLONGIRKRVKFELESPTQWELRAFAVAMHFSLTADRIDDILKVIYDSM 146
QY 339 NHGDARSKRLRELAELTALIKIYVIAQEIKNHLSSTGGINIHDSINIMDKNLGYTD 398
DB 147 NHGDARSKRLRELAELTALIKIYVIAQEIKNHLSSTGGINIHDSINIMDKNLGYTD 206
QY 399 EEIFKASAEKIKLEKMPOTTIOYDGSEKKIVSIKDFLGSEKRTGALGNKNSYKNON 458
DB 207 EEIFKASAEKIKLEKMPOTTIOYDGSEKKIVSIKDFLGSEKRTGALGNKNSYKNON 266
QY 459 NELSHFATTCSDKSRPLNDLVSOQTQLSDITSRNSAIEALNRFIOKDYVQRLDDT 518
DB 267 NELSHFATTCSDKSRPLNDLVSOQTQLSDITSRNSAIEALNRFIOKDYVQRLDDT 326
QY 519 SGR 521
DB 327 SGR 329

RESULT 3
US-08-913-477-17
Sequence 17, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tilball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-17

Query Match 46.6%; Score 243; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.2e-228;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHYDNLONGIRKRVKFELESPTQWELRAFAVAMHFSLTADRIDDILKVIYDSM 338
DB 238 ILKGGHYDNLONGIRKRVKFELESPTQWELRAFAVAMHFSLTADRIDDILKVIYDSM 297
QY 339 NHGDARSKRLRELAELTALIKIYVIAQEIKNHLSSTGGINIHDSINIMDKNLGYTD 398
DB 298 NHGDARSKRLRELAELTALIKIYVIAQEIKNHLSSTGGINIHDSINIMDKNLGYTD 357
QY 399 EEIFKASAEKIKLEKMPOTTIOYDGSEKKIVSIKDFLGSEKRTGALGNKNSYKNON 458
DB 358 EEIFKASAEKIKLEKMPOTTIOYDGSEKKIVSIKDFLGSEKRTGALGNKNSYKNON 417
QY 459 NELSHFATTCSDKSRPLNDLVSOQTQLSDITSRNSAIEALNRFIOKDYVQRLDDT 518
DB 418 NELSHFATTCSDKSRPLNDLVSOQTQLSDITSRNSAIEALNRFIOKDYVQRLDDT 477
QY 519 SGR 521
DB 478 SGR 480

RESULT 4
US-08-913-477-23
Sequence 23, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tilball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-23

Query Match 46.6%; Score 243; DB 2; Length 501;

Best Local Similarity 100.0%; Pred. No. 1.2e-228; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHYNDQNGIKRYKEFLSSPNTQWELRAFMAVHFSLTADRIDDDILKYIVDSM 338
DB 259 ILKGGHYNDQNGIKRYKEFLSSPNTQWELRAFMAVHFSLTADRIDDDILKYIVDSM 318
QY 339 NHHGDAKRLRELAELAEIKIYVIOAEIKHLSSTGTTIHDKSINLMDKNLYGYTD 398
DB 319 NHHGDAKRLRELAELAEIKIYVIOAEIKHLSSTGTTIHDKSINLMDKNLYGYTD 378
QY 399 EEIFASAEYKLEKMPOTTIOVDSSEKKIYIKDFLSEKRTGALGTLKSYSYNKN 458
DB 379 EEIFASAEYKLEKMPOTTIOVDSSEKKIYIKDFLSEKRTGALGTLKSYSYNKN 438
QY 459 NELSHFATCSDKSRPLNDLVYSQKTTQSLDTSRFSNAIEALNRFIOKXDSYQRLDDT 518
DB 439 NELSHFATCSDKSRPLNDLVYSQKTTQSLDTSRFSNAIEALNRFIOKXDSYQRLDDT 498
QY 519 SGK 521
DB 499 SGK 501

RESULT 5
US-08-913-477-21

Sequence 21, Application US/08913477

Patent No. 5985285

GENERAL INFORMATION:

APPLICANT: Tibbali, Richard W.

APPLICANT: Williamson, Ethel D.

APPLICANT: Leary, Sophie E.C.

APPLICANT: Oyston, Petra C.F.

APPLICANT: Bennett, Alice M.

TITLE OF INVENTION: VACCINES FOR PLAGUE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHIE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/913,477

FILING DATE: 15-SEP-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB96/00571

FILING DATE: 13-MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9505059.7

FILING DATE: 13-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9518946.0

FILING DATE: 15-SEP-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9524825.8

FILING DATE: 05-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Crawford, Arthur R.

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 124-599

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4100

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 170 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-913-477-21

Query Match 25.1%; Score 131; DB 2; Length 170;

Best Local Similarity 100.0%; Pred. No. 5.8e-120; Indels 0; Gaps 0;

Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAIALFGIARANAAADLASTATATATVETPARITLTYKEGAPITIMONGNIDT 83
DB 1 MKKISSVIAIALFGIARANAAADLASTATATATVETPARITLTYKEGAPITIMONGNIDT 60
QY 84 ELVGTLLGKYKGTSTVNFDTAAGDPMYLFTSODGNNHOFTRKVIKSDSRDIS 143
DB 61 ELVGTLLGKYKGTSTVNFDTAAGDPMYLFTSODGNNHOFTRKVIKSDSRDIS 120
QY 144 PRVNGENLVGD 154
DB 121 PRVNGENLVGD 131

RESULT 6
US-08-913-477-11

Sequence 11, Application US/08913477

Patent No. 5985285

GENERAL INFORMATION:

APPLICANT: Tibbali, Richard W.

APPLICANT: Williamson, Ethel D.

APPLICANT: Leary, Sophie E.C.

APPLICANT: Oyston, Petra C.F.

APPLICANT: Bennett, Alice M.

TITLE OF INVENTION: VACCINES FOR PLAGUE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYTE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-11

Query Match 24.6%; Score 128; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 4.3e-117;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 TYEGAPITMDNGNITELVGLTIGYKGTSTSVNFTDAAGDPMTLTFTSQDGN 125
DB 24 TYEGAPITMDNGNITELVGLTIGYKGTSTSVNFTDAAGDPMTLTFTSQDGN 83
QY 126 HOFITKVIKGSDFDISPVRNGENLVGDVYLATGSODFFVRISGSGKLAAGXTDA 105
DB 84 HOFITKVIKGSDFDISPVRNGENLVGDVYLATGSODFFVRISGSGKLAAGXTDA 143
QY 186 VTVVSNQ 193
DB 144 VTVVSNQ 151

RESULT 7
US-08-780-496-8
Sequence 8, Application US/08780496
Patent No. 6046048
GENERAL INFORMATION:
APPLICANT: Avi Ashkenazi, Anan Chuntarapai, Kyung Jin Kim
TITLE OF INVENTION: Apo-2 Ligand
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California

COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/780,496
FILING DATE: 08-Jan-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maerschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P0978P1
TELEPHONE: 415/225-5416
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-08-780-496-8

Query Match 4.6%; Score 24; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNHHHHHHSSGHIDDDKHM 24
DB 1 MGNHHHHHHSSGHIDDDKHM 24

RESULT 8
US-08-651-818A-21
Sequence 21, Application US/08651818A
Patent No. 594889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
APPLICANT: Hale, Cynthia A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid

STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-21

Query Match 3.3%; Score 17; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 9,4e-10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 MGNHHHHHHSSGHI 17
|||||
Db 1 MGNHHHHHHSSGHI 17

RESULT 9
US-08-480-604A-24
Sequence 24, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid

STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-604A-24

Query Match 3.3%; Score 17; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 MGNHHHHHHSSGHI 17
|||||
Db 1 MGNHHHHHHSSGHI 17

RESULT 10
US-08-405-496A-24
Sequence 24, Application US/08405496A
Patent No. 5913665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
PREVENTION OF NEUROTOXIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-405-496A-24

Query Match 3.3%; Score 17; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 11

US-08-651-818A-19
Sequence 19, Application US/08651818A
Patent No. 5948889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-19

Query Match 3.3%; Score 17; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 12

US-08-651-818A-23
Sequence 23, Application US/08651818A
Patent No. 5948889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America

ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-23

Query Match 3.3%; Score 17; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 13

US-08-591-989-5
Sequence 5, Application US/08591989
Patent No. 5785721
GENERAL INFORMATION:
APPLICANT: Ross S. Rabin, Sumedha Jayasena
APPLICANT: and Larry Gold
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC
ACID LIGANDS OF ICP4
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.40 MB
MEDIUM TYPE: storage
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,989
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX 49
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 323

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-591-989-5

Query Match 3.3%; Score 17; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGNHHHHHHSSGHI 17
Db 17 MGNHHHHHHSSGHI 33

RESULT 14
US-08-480-604A-26
Sequence 26, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
NUMBER OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-604A-26

Query Match 3.3%; Score 17; DB 1; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGNHHHHHHSSGHI 17
Db 1 MGNHHHHHHSSGHI 17

RESULT 15
US-08-405-496A-26
Sequence 26, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
NUMBER OF INVENTION: NEUTROXIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-405-496A-26

Query Match 3.3%; Score 17; DB 2; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGNHHHHHHSSGHI 17
|||||

Wed Aug 23 11:46:14 2000

us-08-699-716a-2.oligo.ra1

Page 15

Db 1 MGHNNNNNNSSGRT 17

Search completed: August 22, 2000, 17:42:24
Job time: 1124 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:24:36 ; Search time 39.13 Seconds

(Without alignments)
823.970 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGHNNHHHHSSGHIDD.....RFLQKYSVMQRLDDTSGR 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 178050 seqs, 61884766 residues

Word size : 0

Total number of hits satisfying chosen parameters: 178050

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : PIR.64:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243	46.6	326	2	T43594 YOP secretion and
2	191	36.7	326	2	B33601 1crv protein - yer
3	170	32.6	170	2	S13008 capsular antigen F
4	82	15.7	326	2	B37314 regulatory protein
5	17	3.3	168	2	I58315 Wtl - human
6	11	2.1	302	2	A55641 homeotic protein G
7	11	2.1	303	2	A56837 homeotic protein M
8	11	2.1	303	2	B49122 growth arrest-spec
9	11	2.1	303	2	A48130 homeobox protein M
10	11	2.1	354	2	S39406 homeotic protein o
11	11	2.1	355	2	S35345 otx1 protein - mou
12	11	2.1	156541	2	I56541 homeodomain protei
13	11	2.1	414	2	A40350 transcription repr
14	11	2.1	414	2	A48273 delta/ryl/NF-E1/0c
15	11	2.1	414	2	A56418 transcription fact
16	11	2.1	420	2	T39712 hypothetical prote
17	11	2.1	427	2	T42516 hypothetical prote
18	11	2.1	1180	2	S69205 stripe a/b protein
19	10	1.9	132	2	I39193 gene HOXA1 protein
20	10	1.9	133	2	B30242 stem cell protein
21	10	1.9	222	2	I39192 gene HOXA1 protein
22	10	1.9	236	2	S41512 Brn-3b protein - m
23	10	1.9	264	2	S34121 transcription fact
24	10	1.9	298	2	S41469 homeotic protein M
25	10	1.9	309	2	T29293 hypothetical protei
26	10	1.9	321	2	T02987 myb-related protei
27	10	1.9	330	2	I39194 gene HOXA1 protein
28	10	1.9	331	2	A30242 homeotic protein E
29	10	1.9	335	2	G01448 homeobox protein H

30	10	1.9	351	1	K62GHL
31	10	1.9	375	2	A46390
32	10	1.9	387	2	A47446
33	10	1.9	390	1	A38565
34	10	1.9	410	2	I38502
35	10	1.9	411	2	I38502
36	10	1.9	430	2	I58156
37	10	1.9	451	2	S66671
38	10	1.9	469	2	A55909
39	10	1.9	475	2	I37451
40	10	1.9	480	2	A54743
41	10	1.9	580	2	T46024
42	10	1.9	623	2	S33167
43	10	1.9	625	2	S71930
44	10	1.9	628	2	JC2493
45	10	1.9	633	1	A26030
46	10	1.9	657	2	A29454
47	10	1.9	665	2	S02358
48	10	1.9	717	2	S38177
49	10	1.9	754	2	JC4898
50	10	1.9	786	2	A48456
51	10	1.9	950	2	S27473
52	10	1.9	1166	2	T13958
53	10	1.9	1196	2	T14108
54	10	1.9	1212	2	T13804
55	10	1.9	1249	2	T14270
56	10	1.9	1252	2	S36016
57	10	1.9	1273	2	T00338
58	10	1.9	1291	2	T13389
59	10	1.9	1293	2	T14259
60	10	1.9	1305	2	A40879
61	10	1.9	1312	1	B40879
62	10	1.9	1356	2	S51389
63	10	1.9	2212	2	A41098
64	9	1.7	143	1	B64421
65	9	1.7	192	2	T39367
66	9	1.7	255	2	S41511
67	9	1.7	270	2	A26480
68	9	1.7	307	2	A45581
69	9	1.7	311	2	A56235
70	9	1.7	331	2	S78452
71	9	1.7	332	2	JH0465
72	9	1.7	348	2	T04618
73	9	1.7	359	2	S14283
74	9	1.7	420	2	A49642
75	9	1.7	420	2	I59234
76	9	1.7	473	2	A54494
77	9	1.7	568	2	S15008
78	9	1.7	634	2	A54495
79	9	1.7	634	2	A28412
80	9	1.7	654	2	B71633
81	9	1.7	794	2	T39171
82	9	1.7	1116	2	S57382
83	9	1.7	1560	2	T00080
84	9	1.7	2254	2	T09053
85	9	1.7	2273	2	I46478
86	9	1.7	2273	2	I46478
87	9	1.7	2424	2	I46478
88	9	1.7	2424	2	I46478
89	9	1.5	60	2	C64698
90	9	1.5	83	2	T16435
91	8	1.5	102	2	T30119
92	8	1.5	115	2	H22583
93	8	1.5	140	2	A54523
94	8	1.5	183	2	E71524
95	8	1.5	208	2	E71524
96	8	1.5	216	2	E71524
97	8	1.5	250	2	T20257
98	8	1.5	292	2	I51171
99	8	1.5	323	2	I49529
100	8	1.5	323	2	I51751
101	8	1.5	351	2	T20270
102	8	1.5	386	2	S58755

histidine-rich gly
cAMP receptor subu
HNF-3/forx head fa
polycomb (Pc) prot
gene Brn-3b protei
Brn-3.2 - mouse
neuron-derived rec
transforming prote
HBF-G2 (HBF-2) pro
transcription fact
brain factor 1 pro
hypothetical prote
gene pointed prote
neuron-derived rec
serine/threonine-s
knob-associated hi
lamin A - Atrial
SEV7 protein homol
Down-syndrome-crit
occyet wall protei
URB51 protein - sm
SYNGAP-B1 protein
SH3-containing pro
shs protein - fruit
Ras-GTPase activat
occyet wall protei
hypothetical prote
hypothetical prote
phospholipase C (E
phospholipase C (E
ROM2 protein - yea
calcium channel pr
conserved hypothet
hypothetical prote
Brn-3a protein - m
knob protein - pla
Distal-less homeob
transcription acti
POU-domain protein
homeotic protein T
heat shock protein
transcription fact
transcription fact
octamer binding tr
knob-associated hi
gene disco protein
knob protein precu
histidine-rich pro
knob-associated hi
probable amine oxi
hypothetical prote
hypothetical prote
low voltage-activa
calcium channel BI
calcium channel BI
calcium channel BI
calcium channel BI
probable histidine
hypothetical prote
hypothetical prote
hypothetical prote
histidine-rich pro
hypothetical prote
eHAND - mouse
hypothetical prote
transcription fact
transcription fact
homeotic protein o
hypothetical prote
ribosomal protein

103	8	1.5	401	2	S53405	probable membrane	176	7	1.3	321	2	T02936	carbonate dehydrat
104	8	1.5	409	2	S60988	hypothetical prote	177	7	1.3	328	2	S45998	hypothetical prote
105	8	1.5	471	2	T33997	hypothetical prote	178	7	1.3	328	2	G02469	homeotic protein D
106	8	1.5	476	2	S57963	methyl Cpg binding	179	7	1.3	329	1	S61884	carbonate dehydrat
107	8	1.5	495	1	S31223	transcription fact	180	7	1.3	330	1	S61883	carbonate dehydrat
108	8	1.5	499	2	S09880	hypothetical prote	181	7	1.3	330	1	S61882	carbonate dehydrat
109	8	1.5	508	2	S59870	fork head domain p	182	7	1.3	331	1	S48675	carbonate dehydrat
110	8	1.5	559	2	T26141	hypothetical prote	183	7	1.3	338	2	E75576	oxidoreductase - D
111	8	1.5	576	2	S69214	deformed epidermal	184	7	1.3	342	2	T06272	oxanesyl-pyrophosp
112	8	1.5	609	2	A49639	odd-paired - fruit	185	7	1.3	343	2	T29547	hypothetical prote
113	8	1.5	661	2	T46364	hypothetical prote	186	7	1.3	348	2	E64148	hypothetical prote
114	8	1.5	812	2	E75338	DNA gyrase, subunl	187	7	1.3	350	2	S76698	hypothetical prote
115	8	1.5	853	2	T46347	hypothetical prote	188	7	1.3	350	2	A30045	hypothetical prote
116	8	1.5	895	2	JC7089	zinc finger bindin	189	7	1.3	359	2	T21705	homeotic protein r
117	8	1.5	972	2	S35521	DNA topoisomerase	190	7	1.3	362	2	I40304	hypothetical prote
118	8	1.5	1133	2	T30302	P-type ATPase - Te	191	7	1.3	364	2	S11981	outer membrane lip
119	8	1.5	1145	2	T33606	hypothetical prote	192	7	1.3	374	2	T11662	probable major pro
120	8	1.5	1172	2	T00065	hypothetical prote	193	7	1.3	376	2	A49077	transcription init
121	8	1.5	1203	2	T28895	hypothetical prote	194	7	1.3	380	2	S14188	carbonate dehydrat
122	8	1.5	1441	2	T13717	CRAG protein - fru	195	7	1.3	385	1	OMRSP	carbonate dehydrat
123	8	1.5	1557	2	T13160	protein CNK - fru	196	7	1.3	391	2	A49645	seleoprotein P pr
124	8	1.5	185	2	T31611	hypothetical prote	197	7	1.3	392	1	A56691	transcription fact
125	8	1.5	1597	1	BVEFSL	sol protein, large	198	7	1.3	394	2	A70979	phosphoribosylglic
126	8	1.5	1597	2	T08428	gene small optic l	199	7	1.3	398	2	G72333	hypothetical prote
127	8	1.5	1929	2	T21559	hypothetical prote	200	7	1.3	401	2	S09626	conserved hypothet
128	8	1.5	2416	2	T13825	adenomatous polyp	201	7	1.3	405	2	F81076	probable transcript
129	8	1.5	5327	2	T13564	microtubule-assoc	202	7	1.3	406	2	I41084	probable transcript
130	7	1.3	16	1	LEFCH	his operon leader	203	7	1.3	409	2	S40740	type I site-specific
131	7	1.3	77	2	T16436	hypothetical prote	204	7	1.3	410	2	S02167	cyclin A homolog -
132	7	1.3	83	2	C64893	gapc-2 protein, tr	205	7	1.3	415	2	T35834	probable transcript
133	7	1.3	86	2	T34975	S-100 protein beta	206	7	1.3	418	2	D75283	hypothetical prote
134	7	1.3	91	1	BC901B	S-100 protein beta	207	7	1.3	425	1	JH0710	hypothetical prote
135	7	1.3	92	1	BC901B	S-100 protein beta	208	7	1.3	434	1	I51436	transcription fact
136	7	1.3	92	2	A26515	S-100 protein beta	209	7	1.3	434	2	C71434	probable phosphati
137	7	1.3	104	1	G43258	S-100 protein beta	210	7	1.3	439	2	T46375	hypothetical prote
138	7	1.3	105	1	A23696	phosphotransferase	211	7	1.3	444	2	T09474	hypothetical prote
139	7	1.3	111	2	D53589	hypothetical prote	212	7	1.3	449	2	T10650	probable NADH dehy
140	7	1.3	111	2	D53589	hypothetical prote	213	7	1.3	456	2	UC4089	probable NADH dehy
141	7	1.3	117	2	S37150	asir2 protein - tom	214	7	1.3	458	2	S57914	3alpha,7alpha,12al
142	7	1.3	124	2	S42606	hypothetical prote	215	7	1.3	459	2	F75049	probable sakacin P
143	7	1.3	127	2	T44266	ribosomal protein	216	7	1.3	466	2	T44746	hypothetical prote
144	7	1.3	136	2	T22959	hypothetical prote	217	7	1.3	470	2	T17686	probable NADH dehy
145	7	1.3	139	2	T33968	hypothetical prote	218	7	1.3	483	2	T35715	cationic amino aci
146	7	1.3	142	2	T35319	probable regulator	219	7	1.3	490	2	S52830	export protein - S
147	7	1.3	147	2	D64331	hypothetical prote	220	7	1.3	492	2	A41907	HMS1 protein - yea
148	7	1.3	153	1	APBP21	endopeptidase (EC	221	7	1.3	513	1	S60303	methy1-Cpg-binding
149	7	1.3	162	2	S22653	ribonuclease T2 (E	222	7	1.3	513	1	S60304	serine/threonine-s
150	7	1.3	182	2	S51002	ribonuclease precu	223	7	1.3	515	2	T38946	serine/threonine-s
151	7	1.3	182	2	A53803	dihydrofolate redu	224	7	1.3	518	2	S44637	serine-threonine P
152	7	1.3	182	2	H75176	probable dmp deam	225	7	1.3	519	2	A49367	transducin homolog
153	7	1.3	182	2	E71037	probable deoxycyti	226	7	1.3	529	2	T00677	hypothetical prote
154	7	1.3	187	2	T05071	hypothetical prote	227	7	1.3	532	2	S36554	LI protein (altern
155	7	1.3	193	2	H64174	hypothetical prote	228	7	1.3	532	2	T35119	probable aminotran
156	7	1.3	209	2	T17825	hypothetical prote	229	7	1.3	535	1	A39194	periplasmic dipept
157	7	1.3	212	2	S41581	hypothetical prote	230	7	1.3	549	2	T11688	phosphinositide-s
158	7	1.3	236	2	S18783	exonoxin type A pr	231	7	1.3	550	2	T23760	hypothetical prote
159	7	1.3	244	2	T07830	aluminum-induced p	232	7	1.3	564	2	S07744	hypothetical prote
160	7	1.3	252	2	S05629	cytochrome-c oxida	233	7	1.3	570	2	S07744	NADH dehydrogenase
161	7	1.3	254	1	G64045	3-deoxy-manno-occu	234	7	1.3	579	2	JH0820	160K goldi antigen
162	7	1.3	255	2	T24883	hypothetical prote	235	7	1.3	587	2	T16617	hypothetical prote
163	7	1.3	259	2	A70359	hydrogenase expres	236	7	1.3	587	2	S12805	hypothetical prote
164	7	1.3	263	2	T16329	hypothetical prote	237	7	1.3	587	2	S41409	envelysin (EC 3.4.
165	7	1.3	264	2	T02886	carbonate dehydrat	238	7	1.3	601	2	T00119	probable transcript
166	7	1.3	273	2	T33424	hypothetical prote	239	7	1.3	604	2	S56027	hypothetical prote
167	7	1.3	281	2	E70388	conserved hypothet	240	7	1.3	615	2	S54468	hypothetical prote
168	7	1.3	284	2	T35723	cobalt transport s	241	7	1.3	617	2	S54468	AIPI protein - yea
169	7	1.3	290	2	T21868	hypothetical prote	242	7	1.3	620	2	S52494	5-methyltetrahydro
170	7	1.3	294	2	E69759	hypothetical prote	243	7	1.3	622	2	S71342	protein kinase hom
171	7	1.3	297	2	A45335	cytochrome-c oxida	244	7	1.3	627	2	T37293	choleline O-acetyltr
172	7	1.3	306	2	T09067	extensin-like prot	245	7	1.3	647	2	T39141	hypothetical prote
173	7	1.3	314	2	T29363	hypothetical prote	246	7	1.3	682	2	A44493	serum-inducible ki
174	7	1.3	314	2	E64580	catalase-like prot	247	7	1.3	688	2	T23108	hypothetical prote
175	7	1.3	314	2	E71931	hypothetical prote	248	7	1.3	701	2	T20892	hypothetical prote

249	7	1.3	732	2	547073	finger protein HZF
250	7	1.3	733	2	H69411	cell division gont
251	7	1.3	736	2	H75460	conserved hypotnet
252	7	1.3	746	2	G65100	formate C-acetyltr
253	7	1.3	766	2	G71437	probable resistant
254	7	1.3	778	2	T44761	probable preprotel
255	7	1.3	785	2	S57702	hypothetical prote
256	7	1.3	794	2	F72202	hypothetical prote
257	7	1.3	807	2	T42924	glycoprotein B - a
258	7	1.3	808	2	F07020	probable transloca
259	7	1.3	808	2	J02205	Ula7h protein - Ma
260	7	1.3	845	2	T17291	hypothetical prote
261	7	1.3	886	2	JC5085	replication licens
262	7	1.3	892	1	S42228	replication licens
263	7	1.3	892	2	S76350	hypothetical prote
264	7	1.3	901	2	JC7111	tetratricopeptide
265	7	1.3	904	1	T10067	replication licens
266	7	1.3	906	2	RNBPK1	DNA-directed RNA p
267	7	1.3	906	2	A71438	probable resistant
268	7	1.3	906	2	A43817	transfoming prote
269	7	1.3	946	2	C70839	probable mmp13 pro
270	7	1.3	946	2	S48255	probable membrane
271	7	1.3	946	2	S48433	inositol-1,4,5-tri
272	7	1.3	1001	2	T28897	hypothetical prote
273	7	1.3	1133	2	T13384	hypothetical prote
274	7	1.3	1133	2	T13246	lethal(1)1b1 prote
275	7	1.3	1150	2	T13824	LK6 protein kinase
276	7	1.3	1156	2	T18042	ice nucleation pro
277	7	1.3	1289	2	S67200	hypothetical prote
278	7	1.3	1314	2	G02870	KIAA0197 protein -
279	7	1.3	1356	2	T18521	beta transducin-11
280	7	1.3	1361	2	T29435	hypothetical prote
281	7	1.3	1465	2	T23056	hypothetical prote
282	7	1.3	1531	2	H71468	probable outer mem
283	7	1.3	1691	1	D54689	protein-tyrosine-p
284	7	1.3	1705	2	S51672	adenylate cyclase
285	7	1.3	1706	1	OBRC	cyclolysin - Borde
286	7	1.3	1894	2	C54689	protein-tyrosine-p
287	7	1.3	1912	2	A56178	protein-tyrosine-p
288	7	1.3	1979	2	C71622	hypothetical prote
289	7	1.3	2117	2	T36180	CDA peptide synthe
290	7	1.3	2226	2	B47447	calcium channel pr
291	7	1.3	2467	2	D71437	probable resistant
292	7	1.3	3104	2	S20473	fatly-acid synthas
293	7	1.3	3345	2	T13423	hypothetical prote
294	7	1.3	3413	3	T17467	rifamycin polyketi
295	7	1.2	8	2	A61328	trypsin (EC 3.4.21
296	6	1.2	30	2	S15678	acetylcholinestera
297	6	1.2	31	2	S20491	hypothetical prote
298	6	1.2	31	2	T07276	photosystem II pro
299	6	1.2	34	2	A39965	macrocclin-O-methyl
300	6	1.2	35	2	A05057	hypothetical prote
301	6	1.2	53	2	A56512	nuclear pore prote
302	6	1.2	54	2	A31440	ovomucoid, third d
303	6	1.2	54	2	C31436	ovomucoid, third d
304	6	1.2	54	2	A61587	histidine and glut
305	6	1.2	57	2	H64698	hypothetical prote
306	6	1.2	62	2	B69788	hypothetical prote
307	6	1.2	64	2	S57787	hypothetical prote
308	6	1.2	65	2	S01293	acetylcholinestera
309	6	1.2	65	2	S48113	transcription fact
310	6	1.2	65	2	T45078	hypothetical prote
311	6	1.2	65	2	E75022	ssu ribosomal prot
312	6	1.2	66	2	T35489	hypothetical prote
313	6	1.2	66	2	T00821	hypothetical prote
314	6	1.2	70	2	E70985	hypothetical prote
315	6	1.2	71	2	A39566	H+-transporting At
316	6	1.2	76	2	T06117	hypothetical prote
317	6	1.2	77	2	D71821	probable histidine
318	6	1.2	77	2	C75495	hypothetical prote
319	6	1.2	78	2	T21426	hypothetical prote
320	6	1.2	79	2	B72223	hypothetical prote
321	6	1.2	82	2	S22306	hypothetical prote
322	6	1.2	83	2	T48283	gene CC2 protein -
323	6	1.2	83	2	C73083	hypothetical prote
324	6	1.2	85	2	S24453	hypothetical prote
325	6	1.2	87	2	T25505	hypothetical prote
326	6	1.2	89	1	TKMYR	tuberculin-active
327	6	1.2	89	2	C72286	transcription regu
328	6	1.2	89	2	T14959	hypothetical prote
329	6	1.2	89	2	S77075	hypothetical prote
330	6	1.2	91	2	T41989	hypothetical prote
331	6	1.2	92	1	Z1BPC2	regulatory protein
332	6	1.2	92	2	B71637	hypothetical prote
333	6	1.2	92	2	T34146	hypothetical prote
334	6	1.2	92	2	T20721	hypothetical prote
335	6	1.2	93	2	T31266	ferredoxin (Fe-2S
336	6	1.2	95	2	S11788	nolp protein - Rhl
337	6	1.2	99	1	MGRB82	beta-2-microglobul
338	6	1.2	100	2	B71567	hypothetical prote
339	6	1.2	100	2	G81778	conserved hypotnet
340	6	1.2	101	2	B44971	hypothetical prote
341	6	1.2	103	1	MQSA3L	phosphotransferase
342	6	1.2	106	2	A03860	hypothetical prote
343	6	1.2	108	2	S34940	fimbrial protein S
344	6	1.2	108	2	S34942	fimbrial protein S
345	6	1.2	108	2	A45649	microfilament surf
346	6	1.2	109	2	G71077	hypothetical prote
347	6	1.2	111	2	S07367	hypothetical prote
348	6	1.2	112	1	ASLJGH	vpu protein - huma
349	6	1.2	112	2	S12155	hypothetical 13.2K
350	6	1.2	112	2	S03221	conserved hypotnet
351	6	1.2	112	2	H69537	photosystem I 8.4K
352	6	1.2	113	2	S06684	B27R protein - vac
353	6	1.2	113	2	H42528	hypothetical prote
354	6	1.2	115	2	F75140	MBH dehydrogenase
355	6	1.2	117	2	T07281	hypothetical prote
356	6	1.2	118	2	T20862	diacylglycerol kin
357	6	1.2	118	2	H81443	probable chemotaxi
358	6	1.2	120	2	E71160	alpha-lactalbumin
359	6	1.2	120	2	F75002	probable diacygly
360	6	1.2	121	1	LAKGAW	auxin-induced prot
361	6	1.2	122	1	KRECDG	general stress pro
362	6	1.2	122	2	T01558	hypothetical prote
363	6	1.2	123	1	S26183	hypothetical 14.2K
364	6	1.2	125	2	S24831	hypothetical class I
365	6	1.2	127	2	A72388	conserved hypotnet
366	6	1.2	127	2	E65098	matrix protein M2
367	6	1.2	130	2	X14983	biopolymer transpo
368	6	1.2	131	1	C70430	acyl carrier prote
369	6	1.2	132	1	MEFV2J	hypothetical 2b pr
370	6	1.2	133	2	A71855	gene 4.7 protein -
371	6	1.2	133	2	A71855	hypothetical prote
372	6	1.2	134	2	A26860	hypothetical prote
373	6	1.2	134	2	S20499	WD repeat protein
374	6	1.2	134	2	B71392	H33 protein - shie
375	6	1.2	135	1	WABP77	alpha-lactalbumin
376	6	1.2	137	2	S38478	probable profilin
377	6	1.2	137	2	F75043	hypothetical prote
378	6	1.2	139	3	T41428	probable profilin
379	6	1.2	140	1	MMVZM3	ribosomal protein
380	6	1.2	140	2	A60394	hypothetical prote
381	6	1.2	140	2	T06554	ribosomal protein
382	6	1.2	140	2	T19083	hypothetical prote
383	6	1.2	141	2	T06553	tail fiber protein
384	6	1.2	141	2	R3178	transcription regu
385	6	1.2	142	1	F70234	hypothetical prote
386	6	1.2	143	2	T14177	conserved hypotnet
387	6	1.2	143	2	T16332	hypothetical prote
388	6	1.2	144	1	TPBPPI	hypothetical prote
389	6	1.2	144	2	T11756	conserved hypotnet
390	6	1.2	144	2	T42287	hypothetical prote
391	6	1.2	144	2	A69971	hypothetical prote
392	6	1.2	144	2	T25028	hypothetical prote
393	6	1.2	146	2	F64449	plastid-lipid-asso
394	6	1.2	146	2	T07135	

395	6	1.2	147	2	T46623	rearranged T-cell	468	6	1.2	190	2	G69339	conserved hypother
396	6	1.2	147	2	G72384	conserved hypother	469	6	1.2	191	2	S43178	gene s122 protein
397	6	1.2	148	2	T16177	hypothetical prote	470	6	1.2	191	2	T34278	hypothetical prote
398	6	1.2	148	2	JN0568	single-stranded m	471	6	1.2	192	2	JN0577	hypothetical prote
399	6	1.2	148	2	C75126	hypothetical prote	472	6	1.2	192	2	F70222	p-aminobenzoic aci
400	6	1.2	149	2	H72344	conserved hypother	473	6	1.2	193	2	T36959	conserved hypother
401	6	1.2	149	2	F70599	hypothetical prote	474	6	1.2	193	2	C57256	hypothetical prote
402	6	1.2	150	2	T23641	hypothetical prote	475	6	1.2	193	2	E7544	hypothetical prote
403	6	1.2	150	2	T12547	hypothetical prote	476	6	1.2	194	2	D69486	uag-related protei
404	6	1.2	151	2	JN0569	single-stranded m	477	6	1.2	194	2	C44805	eggshell protein -
405	6	1.2	151	2	C41700	Cl protein - rabbi	478	6	1.2	195	2	I46272	trophoblast interf
406	6	1.2	151	2	C75157	protein tyrosine p	479	6	1.2	195	2	F64842	hypothetical prote
407	6	1.2	153	2	G65108	transcription elon	480	6	1.2	197	1	A65081	hypothetical prote
408	6	1.2	154	2	H69320	hypothetical prote	481	6	1.2	197	2	S25825	hypothetical prote
409	6	1.2	154	2	T13439	hypothetical prote	482	6	1.2	197	2	H71003	polysaccharuronase
410	6	1.2	157	2	H69203	conserved hypother	483	6	1.2	198	2	S73275	hypothetical prote
411	6	1.2	158	2	F70955	hypothetical prote	484	6	1.2	198	2	B71214	hypothetical prote
412	6	1.2	158	2	T05097	hypothetical prote	485	6	1.2	198	2	S24335	hypothetical prote
413	6	1.2	158	2	D75069	hypothetical prote	486	6	1.2	198	2	F55578	H+-transporting Ar
414	6	1.2	158	2	B71096	hypothetical prote	487	6	1.2	199	2	C75363	hypothetical prote
415	6	1.2	159	2	S72326	glucan 1,3-beta-gl	488	6	1.2	199	2	T05283	hydrolase family p
416	6	1.2	159	2	S33585	lipoprotein smpa p	489	6	1.2	199	2	S59741	hypothetical prote
417	6	1.2	161	2	B81357	probable lipoprote	490	6	1.2	200	2	S75359	PAC10 protein - ye
418	6	1.2	161	2	E43719	uree protein - Pro	491	6	1.2	200	2	T36125	hypothetical prote
419	6	1.2	162	2	A27398	allophycocyanin al	492	6	1.2	201	2	T38713	hypothetical prote
420	6	1.2	162	2	T45343	hypothetical prote	493	6	1.2	202	2	S74224	hypothetical prote
421	6	1.2	162	2	H70927	probable rpsp prot	494	6	1.2	202	2	T16495	hypothetical prote
422	6	1.2	162	2	T33127	hypothetical prote	495	6	1.2	203	2	S67607	hypothetical prote
423	6	1.2	163	2	I40858	superoxide dismuta	496	6	1.2	203	2	T33126	probable membrane
424	6	1.2	163	2	S28136	gas vesicle protei	497	6	1.2	204	2	B75539	hypothetical prote
425	6	1.2	164	1	zpcrl	lipoprotein signal	498	6	1.2	204	2	T34944	probable phosphogl
426	6	1.2	164	1	T24726	hypothetical prote	499	6	1.2	204	2	S67295	probable lipoprote
427	6	1.2	167	2	B70458	lipoprotein signal	500	6	1.2	204	2	D75154	probable membrane
428	6	1.2	168	2	I40346	ompl6 protein - Br	501	6	1.2	205	2	E69327	hypothetical prote
429	6	1.2	168	2	C71450	hypothetical prote	502	6	1.2	205	2	E75310	ribonuclease HII (
430	6	1.2	168	2	T25336	hypothetical prote	503	6	1.2	206	2	T33064	hypothetical prote
431	6	1.2	168	2	A70479	hypothetical prote	504	6	1.2	206	2	T25384	hypothetical prote
432	6	1.2	169	1	S23478	probable benzozate	505	6	1.2	208	2	T24446	hypothetical prote
433	6	1.2	170	2	B37152	lipoprotein signal	506	6	1.2	208	2	D71258	probable gipg prot
434	6	1.2	172	2	H65026	hypothetical prote	507	6	1.2	209	2	H81149	conserved hypother
435	6	1.2	172	2	T36107	probable serine/ar	508	6	1.2	210	2	S25657	T-cell surface gly
436	6	1.2	172	2	C72007	low calcium respon	509	6	1.2	210	2	S64376	hypothetical prote
437	6	1.2	172	2	G81533	type III secretion	510	6	1.2	210	2	JC5512	hypothetical prote
438	6	1.2	173	1	YOECP	fibinrial protein p	511	6	1.2	211	2	B71809	SMAR protein 23 -
439	6	1.2	173	2	I41201	E-minor fibinrial p	512	6	1.2	211	2	E64710	hypothetical prote
440	6	1.2	173	2	B27743	pape fibinrial prot	513	6	1.2	211	2	C59091	hypothetical prote
441	6	1.2	173	2	S25210	PRSE protein - Esc	514	6	1.2	211	2	T00501	hypothetical prote
442	6	1.2	173	2	C81549	conserved hypother	515	6	1.2	213	2	E64690	RBRIC protein homo
443	6	1.2	174	2	JC1394	pape fibinrial prot	516	6	1.2	213	2	C68641	response regulator
444	6	1.2	175	2	C27743	pape fibinrial prot	517	6	1.2	213	2	S75247	ATP phosphoribosyl
445	6	1.2	175	2	C69498	conserved hypother	518	6	1.2	214	2	I49758	hypothetical prote
446	6	1.2	175	2	T46136	remorin-like prote	519	6	1.2	214	2	D64163	hypoxanthine phosph
447	6	1.2	175	2	I53285	glucocorticoid-reg	520	6	1.2	214	2	T47110	3-oxodipate CoA-t
448	6	1.2	175	2	A53523	endocrine secretor	521	6	1.2	215	2	B75088	hypothetical prote
449	6	1.2	175	2	JC5912	lysyl protein - Lac	522	6	1.2	216	2	T26054	hypothetical prote
450	6	1.2	175	2	T13323	protein R175 - Lac	523	6	1.2	218	1	RTMSG	hypoxanthine phosph
451	6	1.2	177	2	D71869	hypothetical prote	524	6	1.2	218	1	RTHYG	hypoxanthine phosph
452	6	1.2	178	2	H81207	conserved hypother	525	6	1.2	218	2	S43043	hypoxanthine (guan
453	6	1.2	179	1	R5BS5	ribosomal protein	526	6	1.2	218	2	S21474	hypoxanthine phosph
454	6	1.2	179	2	S29313	inorganic pyrophos	527	6	1.2	218	2	T51842	hypoxanthine phosph
455	6	1.2	179	2	B75636	transcription regu	528	6	1.2	218	2	S18140	hypoxanthine phosph
456	6	1.2	180	2	F75073	pyrazinamidase/nic	529	6	1.2	218	2	T26826	hypoxanthine phosph
457	6	1.2	181	2	H69502	hypothetical prote	530	6	1.2	219	1	C64625	probable thiamin-p
458	6	1.2	182	2	G70698	probable pp1a prot	531	6	1.2	219	2	T32443	hypothetical prote
459	6	1.2	182	2	A02946	keratin, 59K type	532	6	1.2	220	2	B53312	hypothetical prote
460	6	1.2	182	2	A02947	keratin, 60K type	533	6	1.2	220	2	A44805	deoxyribose-phosph
461	6	1.2	183	2	T27236	hypothetical prote	534	6	1.2	220	2	T44702	eggshell protein -
462	6	1.2	184	2	S14688	hypothetical prote	535	6	1.2	221	2	G73322	glutaredoxin-relat
463	6	1.2	184	2	T06673	hypothetical prote	536	6	1.2	221	2	A70045	pectate lyase homo
464	6	1.2	184	2	T24251	hypothetical prote	537	6	1.2	221	2	T07079	leucine-rich repea
465	6	1.2	184	2	PC4032	nitrite reductase	538	6	1.2	222	2	T29457	probable two-compo
466	6	1.2	186	2	G72373	lipoprotein signal	539	6	1.2	222	2	S34563	gene M1 protein -
467	6	1.2	188	2	S67051	hypothetical prote	540	6	1.2	222	2	B49599	polymerase-associa

541	6	1.2	223	1	QOBE26	BzfF2 protein - hu	614	6	1.2	256	2	D69845	thiamin biosynthes
542	6	1.2	223	2	T07098	catechol oxidase (615	6	1.2	256	2	A69129	conserved hypotet
543	6	1.2	224	1	KICHT	thymidine kinase (616	6	1.2	257	2	A28664	enterotoxin A prec
544	6	1.2	224	2	B70327	ABC transporter -	617	6	1.2	257	2	D71544	hypothetical prote
545	6	1.2	226	2	T10660	photosystem II pro	618	6	1.2	258	1	F69374	conserved hypotet
546	6	1.2	226	1	G64387	hypothetical prote	619	6	1.2	258	2	T01461	carbonate dehydrat
547	6	1.2	226	2	C71307	probable ABC trans	620	6	1.2	258	2	D70758	hypothetical prote
548	6	1.2	226	2	A70565	probable cutinase	621	6	1.2	258	2	B72697	hypothetical prote
549	6	1.2	226	2	G69129	hypothetical prote	622	6	1.2	260	2	B64563	hypothetical prote
550	6	1.2	227	2	T11080	H+-transporting AT	623	6	1.2	261	2	E10321	3ik protein - frog
551	6	1.2	227	2	T46686	pcx protein [limp	624	6	1.2	261	2	F81376	probable integral
552	6	1.2	230	2	T44953	flagella-related p	625	6	1.2	262	2	F72333	hypothetical prote
553	6	1.2	230	2	T18757	hypothetical prote	626	6	1.2	263	2	B17494	probable php hydrol
554	6	1.2	231	1	F1SP3	photosystem I chl	627	6	1.2	263	2	B81653	mtc protein, prob
555	6	1.2	232	2	S77017	sensory transducti	628	6	1.2	264	2	G69632	flagellar hook pro
556	6	1.2	232	2	S75919	hypothetical prote	629	6	1.2	264	2	S66439	hypothetical trans
557	6	1.2	234	2	S22360	hypothetical prote	630	6	1.2	265	1	B65132	hypothetical prote
558	6	1.2	235	2	S60158	chlamyopsin - Chla	631	6	1.2	265	2	F70907	hypothetical prote
559	6	1.2	235	2	H72093	tRNA pseudouridine	632	6	1.2	266	2	T36287	probable integral
560	6	1.2	235	2	D81576	tRNA pseudouridine	633	6	1.2	267	2	S76499	flagellar hook-bas
561	6	1.2	236	2	D69423	branched-chain am	634	6	1.2	268	2	G69623	hypothetical prote
562	6	1.2	237	2	B75214	hypothetical prote	635	6	1.2	268	2	B39429	CAMP response elem
563	6	1.2	237	2	T35855	probable ABC trans	636	6	1.2	268	2	B37279	enhancer-binding p
564	6	1.2	237	2	C40784	hypothetical prote	637	6	1.2	268	2	C37280	C/EBP-related prot
565	6	1.2	239	1	T35279	probable dehydroge	638	6	1.2	268	2	A56446	ig heavy chain V r
566	6	1.2	239	1	B63406	probable 3-isoprop	639	6	1.2	269	2	A40225	transcription acti
567	6	1.2	239	2	T20255	hypothetical prote	640	6	1.2	269	2	A47008	transcription acti
568	6	1.2	239	2	S59050	tryptophan synthas	641	6	1.2	269	2	S51814	hypothetical prote
569	6	1.2	239	2	B72573	hypothetical prote	642	6	1.2	270	2	S75051	lactose transport
570	6	1.2	240	2	T45727	hypothetical prote	643	6	1.2	270	2	T03635	plastid-lipid-asso
571	6	1.2	240	2	B71615	Ygb protein, 1st	644	6	1.2	270	2	JE0167	nitric-oxide reduc
572	6	1.2	241	2	F71478	probable pseudouri	645	6	1.2	271	2	S25006	calretinin - rat
573	6	1.2	241	2	E81741	ribosomal large ch	646	6	1.2	271	2	A60253	transcription fact
574	6	1.2	242	2	D81438	Ercc1-like excisio	647	6	1.2	271	2	JC6553	hypothetical prote
575	6	1.2	242	2	E71621	hypothetical prote	648	6	1.2	271	2	D70591	granaticin polyket
576	6	1.2	242	2	T21226	trans-zeatin secre	649	6	1.2	272	1	T01767	membrane protein 1
577	6	1.2	243	1	JOAG32	probable adenylate	650	6	1.2	272	1	S05397	hypothetical prote
578	6	1.2	243	1	S06738	trans-zeatin secre	651	6	1.2	273	1	MMAGCG	hypothetical prote
579	6	1.2	243	2	S03122	trans-zeatin secre	652	6	1.2	273	2	E64316	hypothetical prote
580	6	1.2	243	2	D71450	hypothetical prote	653	6	1.2	273	2	JC1166	type II site-speci
581	6	1.2	243	2	T12734	structural protein	654	6	1.2	274	1	J00957	myb-related protei
582	6	1.2	244	2	T31838	hypothetical prote	655	6	1.2	274	2	S22308	type II site-speci
583	6	1.2	244	2	A55456	transcription acti	656	6	1.2	274	2	A64978	hypothetical prote
584	6	1.2	245	2	T24565	hypothetical prote	657	6	1.2	275	2	T02334	hypothetical prote
585	6	1.2	246	1	TRDGC	trypsin (EC 3.4.21	658	6	1.2	277	2	B69479	uracil accessory p
586	6	1.2	246	2	I50209	reverse transcript	659	6	1.2	278	2	F72064	conserved hypotet
587	6	1.2	247	2	S53863	H+-transporting AT	660	6	1.2	278	2	T33129	solute binding pro
588	6	1.2	247	2	A43846	lipoprotein lpnpa -	661	6	1.2	278	2	B81601	hypothetical prote
589	6	1.2	247	2	T29733	hypothetical prote	662	6	1.2	279	2	G72637	ABC transporter, p
590	6	1.2	248	2	T40343	hypothetical ubiqu	663	6	1.2	280	2	H72389	hypothetical prote
591	6	1.2	248	2	S64486	Yfp1 protein - yea	664	6	1.2	280	2	S42479	transcription regu
592	6	1.2	248	2	T21786	hypothetical prote	665	6	1.2	281	2	S44230	hypothetical prote
593	6	1.2	249	2	G64876	hypothetical prote	666	6	1.2	282	2	C64435	strf protein - Str
594	6	1.2	249	2	T21785	hypothetical prote	667	6	1.2	283	2	S52928	shikimate 5-dehydr
595	6	1.2	249	2	G75189	hypothetical prote	668	6	1.2	283	2	T31122	XSGU1 protein - Af
596	6	1.2	250	2	T20256	hypothetical prote	669	6	1.2	284	2	G72107	probable transcrip
597	6	1.2	250	2	T20258	hypothetical prote	670	6	1.2	284	2	G71563	oligonucleotide trans
598	6	1.2	250	2	F72057	glutamine binding	671	6	1.2	285	2	S78841	peptide ABC transp
599	6	1.2	250	2	C64304	ABC transporter su	672	6	1.2	285	2	C72516	hypothetical prote
600	6	1.2	250	2	E81609	amino acid ABC tra	673	6	1.2	285	2	T21962	hypothetical prote
601	6	1.2	252	1	S59047	indole-3-glycerol	674	6	1.2	286	2	C64538	cytochrome-c oxida
602	6	1.2	252	1	RPECCD	deoxyribose operon	675	6	1.2	286	2	T10129	thymidylate synth
603	6	1.2	252	2	T40604	hypothetical prote	676	6	1.2	287	2	F70925	probable ribosomal
604	6	1.2	252	2	T24939	hypothetical prote	677	6	1.2	287	2	A47038	prolyl aminopeptid
605	6	1.2	252	2	A32489	oncostatin M precu	678	6	1.2	288	2	B41768	orotidine-5'-phosp
606	6	1.2	253	2	B71346	conserved spool reg	679	6	1.2	288	2	T40472	probable spindle p
607	6	1.2	253	2	F64741	conserved hypotet	680	6	1.2	289	2	C70400	ferredoxin oxidore
608	6	1.2	253	2	T15385	hypothetical prote	681	6	1.2	289	2	A69161	hypothetical prote
609	6	1.2	253	2	B70483	ABC transporter -	682	6	1.2	289	2	S16556	hypothetical prote
610	6	1.2	255	1	S04899	myb-related protei	683	6	1.2	289	2	S13800	centrosomin A - mo
611	6	1.2	255	1	E72360	pyridine-5-carbox	684	6	1.2	290	2	S55276	N-acetyltransferas
612	6	1.2	255	2	A75502	probable acetyltra	685	6	1.2	291	1	EWMTG	gamma-glutadin B pr
613	6	1.2	256	2	S37893	csfA protein homol	686	6	1.2	292	2	C75024	hydrogenase, gamma

833	6	1.2	351	2	E75438	lyptophanyl-tRNA
834	6	1.2	351	2	B72315	conserved hypotnet
835	6	1.2	352	2	JS0023	glycerol-3-phospha
836	6	1.2	352	2	T20729	hypothetical prote
837	6	1.2	352	2	D69123	coenzyme F420-redu
838	6	1.2	352	2	T18794	hypothetical prote
839	6	1.2	353	2	S11790	glycerol-3-phospha
840	6	1.2	353	2	C31751	protein kinase (EC
841	6	1.2	353	2	T35921	probable oxidoredu
842	6	1.2	354	2	T36559	probable epoxide h
843	6	1.2	355	1	ROEBPM	recombination prot
844	6	1.2	355	1	S30315	DNA-methyltransfer
845	6	1.2	355	2	S41686	geranylgeranyltran
846	6	1.2	355	2	S11152	oligopeptide trans
847	6	1.2	355	2	C39725	hypothetical prote
848	6	1.2	356	2	T41764	ACKNPV orf18 - Bom
849	6	1.2	358	2	T46532	protein co-factor
850	6	1.2	359	1	DEUT1C	glycerol-3-ph
851	6	1.2	359	1	DEUT1B	glycerol-3-ph
852	6	1.2	359	2	T32287	hypothetical prote
853	6	1.2	360	2	D69025	pleiotropic regula
854	6	1.2	360	2	S53536	MAP kinase (EC 2.7
855	6	1.2	360	2	T49066	cell division cycl
856	6	1.2	360	2	S12662	S-adenosylmethioni
857	6	1.2	360	2	S28047	TUBB3 protein - po
858	6	1.2	360	2	A72423	D-mannosate hydrol
859	6	1.2	361	2	T01934	adenosylmethionine
860	6	1.2	361	2	S74433	GDP-D-mannose dehy
861	6	1.2	362	2	D81606	conserved hypotnet
862	6	1.2	362	2	A12061	cit74 hypothetical
863	6	1.2	363	2	S23137	glycerol-3-phospha
864	6	1.2	363	2	T26385	hypothetical prote
865	6	1.2	363	2	S16998	polygalacturonase
866	6	1.2	363	2	S49222	adenosylmethionine
867	6	1.2	363	2	S43159	outer membrane por
868	6	1.2	363	2	UC1275	phospho-N-acetylmu
869	6	1.2	363	2	UX0285	sensor protein bas
870	6	1.2	363	2	T27335	hypothetical prote
871	6	1.2	364	2	T02801	probable membrane
872	6	1.2	365	2	B54138	Fc-binding protein
873	6	1.2	365	2	A71005	probable cell divi
874	6	1.2	365	2	F70988	hypothetical prote
875	6	1.2	366	2	S75410	tyrosine--tRNA lig
876	6	1.2	368	2	F72281	hypothetical prote
877	6	1.2	368	2	E71204	hypothetical prote
878	6	1.2	369	1	TYFVAF	transforming prote
879	6	1.2	369	1	B64921	conserved hypotnet
880	6	1.2	369	2	H64611	probable transamin
881	6	1.2	369	2	G71902	variable transamin
882	6	1.2	369	2	S11980	variable major pro
883	6	1.2	369	2	T08266	conserved hypotnet
884	6	1.2	369	2	C75043	hypothetical prote
885	6	1.2	370	2	T57555	c-Ha1 protein - mo
886	6	1.2	370	2	T45773	mucin-like protein
887	6	1.2	371	2	C64499	isocitrate dehydro
888	6	1.2	371	2	T45776	hypothetical protei
889	6	1.2	372	2	H81291	mucin-like protei
890	6	1.2	372	2	S24378	hypothetical prote
891	6	1.2	372	2	C64791	ybak protein - Esc
892	6	1.2	372	2	C70590	probable nict prot
893	6	1.2	373	2	A41918	transforming growt
894	6	1.2	373	2	G64482	hypothetical GTP-b
895	6	1.2	373	2	T10577	hypothetical prote
896	6	1.2	374	2	S76392	phosphoglycerate k
897	6	1.2	374	2	A41096	GTP-binding regula
898	6	1.2	374	2	T05891	myb-related protei
899	6	1.2	374	2	A57583	histone acetyltran
900	6	1.2	374	2	S46810	hypothetical prote
901	6	1.2	374	2	T34435	hypothetical prote
902	6	1.2	374	2	F81693	n1S protein, prob
903	6	1.2	375	2	A64398	hypothetical prote
904	6	1.2	376	2	C81272	probable amlinoran
905	6	1.2	376	2	T43846	translation elonga
906	6	1.2	377	2	I57490	guanine nucleotide
907	6	1.2	377	2	F69935	conserved hypotnet
908	6	1.2	377	2	A35795	cardonnet-cytochr
909	6	1.2	378	1	S17412	ubiquitinol--cytoch
910	6	1.2	378	2	S54056	probable membrane
911	6	1.2	379	2	E70937	hypothetical prote
912	6	1.2	379	2	I40873	serine proteinase
913	6	1.2	379	2	S42543	hypothetical prote
914	6	1.2	379	2	E72284	oxidoreductase, al
915	6	1.2	380	2	H71191	hypothetical prote
916	6	1.2	380	2	E69319	flavoprotein reduc
917	6	1.2	380	2	D44490	retrovirus-related
918	6	1.2	381	1	CMBO	chymosin (EC 3.4.2
919	6	1.2	381	1	CM5HB	chymosin (EC 3.4.2
920	6	1.2	381	2	E75458	acyl-CoA dehydroge
921	6	1.2	382	2	T14596	gag protein - alim
922	6	1.2	382	2	T46707	probable Iprn prot
923	6	1.2	384	2	A70805	conserved hypotnet
924	6	1.2	384	2	T08318	syndecan-3 - chick
925	6	1.2	384	2	A44146	aldehyde dehydroge
926	6	1.2	385	2	JC5019	hypothetical prote
927	6	1.2	385	2	T19201	reverse transcript
928	6	1.2	385	2	S54987	reverse transcript
929	6	1.2	385	2	S54988	reverse transcript
930	6	1.2	385	2	S54990	reverse transcript
931	6	1.2	385	2	S54991	reverse transcript
932	6	1.2	385	2	S54992	reverse transcript
933	6	1.2	385	2	S54995	reverse transcript
934	6	1.2	385	2	S54997	reverse transcript
935	6	1.2	387	2	E75184	hypothetical prote
936	6	1.2	389	2	T22611	hypothetical prote
937	6	1.2	390	1	A42724	transcription init
938	6	1.2	390	1	TVMCB	transforming prote
939	6	1.2	390	2	H70353	onitiline decarbox
940	6	1.2	390	2	I39975	sarcosine oxidase
941	6	1.2	391	2	D64366	hypothetical prote
942	6	1.2	391	2	S17818	bacteriophage phi
943	6	1.2	391	2	B70608	probable PPE prote
944	6	1.2	392	2	T00669	polygalacturonase
945	6	1.2	392	2	T37173	probable cystathio
946	6	1.2	392	2	S11998	finger protein odd
947	6	1.2	393	2	T23602	hypothetical prote
948	6	1.2	393	2	E81691	conserved hypotnet
949	6	1.2	394	2	S57099	hypothetical prote
950	6	1.2	395	2	A47701	hypothetical prote
951	6	1.2	395	2	B64338	aspartic proteinas
952	6	1.2	395	2	S44454	hypothetical prote
953	6	1.2	396	2	PS0075	transcription fact
954	6	1.2	396	2	S34963	mosin heavy chain
955	6	1.2	396	2	S26391	ribx protein - Shi
956	6	1.2	397	2	S30023	tall sheath protei
957	6	1.2	397	2	T00098	antiviral protein
958	6	1.2	398	2	A55581	hypothetical prote
959	6	1.2	398	2	T30013	nasc protein - Kle
960	6	1.2	398	2	S77529	hypothetical prote
961	6	1.2	398	2	T01921	hypothetical prote
962	6	1.2	399	2	T38388	translation elonga
963	6	1.2	400	1	A28172	hypothetical wd-40
964	6	1.2	400	1	A39822	spasmolysin precu
965	6	1.2	400	2	C69757	leukosialin precu
966	6	1.2	400	2	S70585	transporter homolo
967	6	1.2	400	2	S24759	nitrite reductase
968	6	1.2	400	2	A42727	probable transposa
969	6	1.2	400	2	G71900	hypothetical prote
970	6	1.2	401	2	E75213	n1S protein PAB01
971	6	1.2	401	2	H72765	hypothetical prote
972	6	1.2	402	2	A70559	probable flag endo
973	6	1.2	402	2	A70882	hypothetical prote
974	6	1.2	402	2	T45712	probable PPE prote
975	6	1.2	403	2	E81096	NAD-dependent mala
976	6	1.2	403	2	T25453	uracil permease NM
977	6	1.2	404	2	T34085	hypothetical prote
978	6	1.2	404	2	T34085	hypothetical prote

ALIGNMENTS

```

OY      459  NELSFATCTCDKSRPLNDLVSOQKTOLTSDITSRNSAIEALNFEIQKDYDSVMORLLDDT  518
Db      264  NELSFATCTCDKSRPLNDLVSOQKTOLTSDITSRNSAIEALNFEIQKDYDSVMORLLDDT  323

OY      519  SGK  521
          |||
          |||
          |||
Db      324  SGK  326

RESULT  2
B33601
lcerv protein - Yersinia pestis
C:/Species: Yersinia pestis
C:/Date: 17-Jan-1990 #sequence_revision 17-Jan-1990 #text_change 23-Mar-1993
C:/Accession: B33601
R:/Price: S.B.; Leung, K.Y.; Barve, S.S.; Straley, S.C.
J. Bacteriol. 171, 5646-5653, 1989
A:/Title: Molecular analysis of lcrGVH, the v antigen operon of Yersinia pestis.
A:/Reference number: A33601; MUID:90008806
A:/Accession: B33601
A:/Status: Preliminary
A:/Molecule type: DNA
A:/Residues: 1526 <PR1>
A:/Cross-references: GB:M26405

```

Query March 36.7% Score 191; DB 2; Length 326;
 Best Local Similarity 100.0%; Pred. No. 8.1e-183;
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 LKVIYDSNNHHGGDARSKRLRELAELTAELKITYSVIQAIEINKHLSSTGGINIHDSKINIMLD 390
 Db 136 LKVIYDSNNHHGGDARSKRLRELAELTAELKITYSVIQAIEINKHLSSTGGINIHDSKINIMLD 195
 QY 391 KNLGYTDEELFEFKSAEKKILEKMPQTTIYQDGESEKKIVSITKDFLGSENNKRTGALGNLKN 450
 Db 196 KNLGYTDEELFEFKSAEKKILEKMPQTTIYQDGESEKKIVSITKDFLGSENNKRTGALGNLKN 255
 QY 451 SYSYNNKNNELSHPATTCGDSRPLNDVSKOTQOLSDITRFSNAITALNRFTOKYDVS 510
 Db 256 SYSYNNKNNELSHPATTCGDSRPLNDVSKOTQOLSDITRFSNAITALNRFTOKYDVS 315
 QY 511 MQRLLDPTSGK 521
 Db 316 MQRLLDPTSGK 326

RESULT 3
 S13008
 capsular antigen F1 precursor - Yersinia pestis plasmid pWT1
 C:Species: Yersinia pestis
 C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 04-Mar-2000
 C:Accession: S13008; T14702; T15015; S23725
 R:Galayov, E.E.; Smiltsov, O.Y.; Karlishev, A.V.; Volkovoy, K.I.; Deneznyuk, A.I.; Nazim
 FEBS Lett. 277, 230-232, 1990
 A:Title: Nucleotide sequence of the Yersinia pestis gene encoding F1 antigen and the
 A:Reference number: S13008; MUID:91099503
 A:Accession: S13008
 A:Molecule type: DNA
 A:Residues: 1-170 <GAL>
 A:Cross-references: EMBL:X61996; NID:q48620; PIDN:CAA43966.1; PID:q48621
 R:Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Carrea
 submitted to the EMBL data library, March 1998
 A:Description: Structural organization of virulence determinants in three Yersinia pe
 A:Reference number: Z18168
 A:Accession: T14702
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-170 <HUP>
 A:Cross-references: EMBL:AF053947; NID:q2996286; PID:q2996338; PIDN:AA013218.1
 R:Findler, L.E.; Plano, G.V.; Butland, V.; Mayhew, G.F.; Blattner, F.R.
 Infect. Immun. 66, 5731-5742, 1998

A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIM5 plasmid
 A:Reference number: Z18286; MUID:99043898
 A:Accession: F15015
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-170 <LID>
 A:Cross-references: EMBL:AF074611; NID:93883003; PID:93883098; PIDN:AAC82758.1
 C:Genetics:
 A:Gene: cat1
 A:Genome: plasmid PMT1
 C:Superfamily: Yersinia pestis plasmid PMT1 capsular antigen FI
 F:1-21/Domain: signal sequence status predicted <SIG>
 F:22-170/Product: capsular antigen FI status predicted <MAT>

Query Match 32.6%; Score 170; DB 2; Length 170;
 Best Local Similarity 100.0%; Pred. No. 4,5e-162;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAALFGTITANADLTASTATATLVEPARITLYKKGAPITIMDNGNIDT 83
 |||||||
 DB 1 MKKISSVIAALFGTITANADLTASTATATLVEPARITLYKKGAPITIMDNGNIDT 60

QY 84 ELLVGLTLGGKTKTSTSVNFTDAAGPMYITFTSOGNNHOFITKYIGKDSRFDIS 143
 |||||||
 DB 61 ELLVGLTLGGKTKTSTSVNFTDAAGPMYITFTSOGNNHOFITKYIGKDSRFDIS 120

QY 144 PLYNGENLVGDVVLATGSGDFEVRISGSGKGLAGKTYTDAVTYVSNQ 193
 |||||||
 DB 121 PLYNGENLVGDVVLATGSGDFEVRISGSGKGLAGKTYTDAVTYVSNQ 170

RESULT 4
 B37314
 Regulatory protein IcrY - Yersinia pseudotuberculosis
 C:Species: Yersinia pseudotuberculosis
 C:Date: 30-Jan-1993 #sequence_revision 30-Jan-1993 #text_change 08-Oct-1999
 C:Accession: B37314
 R:Berghman, T.; Hakansson, S.; Forsberg, A.; Norlander, L.; Macellaro, A.; Baackman, A.; J. Bacteriol. 173, 1607-1616, 1991
 A:Title: Analysis of the Y antigen IcrYVH-yopBD operon of Yersinia pseudotuberculosis:
 A:Reference number: A37314; MUID:91154114
 A:Accession: B37314
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-326 <BER>
 A:Cross-references: GB:M57893; NID:9155456; PIDN:AAA27645.1; PID:9155458

Query Match 15.7%; Score 82; DB 2; Length 326;
 Best Local Similarity 100.0%; Pred. No. 8,7e-74;
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 MRAYQONQHITIEDLEKRVDELTHGSSVLELVOLYKKNIDISIKYDPKRDSEVA 255
 |||||||
 DB 1 MRAYQONQHITIEDLEKRVDELTHGSSVLELVOLYKKNIDISIKYDPKRDSEVA 60

QY 256 NRVITDDIELKRIAYFLPED 277
 |||||||
 DB 61 NRVITDDIELKRIAYFLPED 82

RESULT 5
 I58315
 WT1 - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
 C:Accession: I58315
 R:Hamilton, T.B.; Barilla, K.C.; Romanuk, P.J.
 Nucleic Acids Res. 23, 277-284, 1995
 A:Title: High affinity binding sites for the 'Wilms' tumour suppressor protein WT1.
 A:Reference number: I58315; MUID:95166649
 A:Accession: I58315

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-168 <RES>
 A:Cross-references: GB:S75264; NID:9896246; PIDN:AB33443.1; PID:9896247
 C:Genetics:
 A:Gene: GDB:WT1
 A:Cross-references: GDB:120496; OMIM:194070
 A:Map position: 11p13-11p13

Query Match 3.3%; Score 17; DB 2; Length 168;
 Best Local Similarity 100.0%; Pred. No. 4,9e-09;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNHHHHHHHSSGHI 17
 |||||||
 DB 1 MGNHHHHHHHSSGHI 17

RESULT 6
 A55641
 homeotic protein GAX - human
 C:Species: Homo sapiens (man)
 C:Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 24-Sep-1999
 C:Accession: A55641
 R:LePage, D.F.; Altomare, D.A.; Testa, J.R.; Walsh, K.
 Genomics 24, 535-540, 1994
 A:Title: Molecular cloning and localization of the human GAX gene to 7p21.
 A:Reference number: A55641; MUID:95229154
 A:Accession: A55641
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-302 <LEP>
 A:Cross-references: GB:I36328; NID:9557549; PIDN:AA58497.1; PID:9557550
 C:Genetics:
 A:Gene: GAX
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:186-242/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 302;
 Best Local Similarity 100.0%; Pred. No. 0.0085;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHHHHHHHHH 12
 |||||||
 DB 67 GHHHHHHHHH 77

RESULT 7
 A56837
 homeotic protein MOX2 - human
 C:Species: Homo sapiens (man)
 C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 24-Sep-1999
 C:Accession: A56837; S52168
 R:Girgionou, M.; Kastlinski, M.C.; Modl, W.; Theodorakis, K.; Mankoo, B.; Pachnis, V.
 Genomics 26, 550-555, 1995
 A:Title: Isolation of the human MOX2 homeobox gene and localization to chromosome 7p2
 A:Reference number: A56837; MUID:95331791
 A:Accession: A56837
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-303 <GR1>
 A:Cross-references: EMBL:X82629; NID:9732790; PIDN:CA57949.1; PID:9732791
 R:Girgionou, M.; Kastlinski, M.C.; Modl, W.; Theodorakis, K.; Mankoo, B.; Pachnis, V.
 submitted to the EMBL Data Library, November 1994
 A:Description: Isolation of the human MOX2 homeobox gene and localization to chromoso
 A:Reference number: S52168
 A:Accession: S52168
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-7, 'R', '9-110, 'R', 112-122, 'Q', 124-157, 'V', 159-303 <GR2>

A:Cross-references: EMBL:X82629
 C:Genetics:
 A:Gene: GDB:MOX2
 A:Cross-references: GDB:120703; OMIM:155970
 A:Map position: 3q13-3q13
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:187-243/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0085;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 Db 67 GHHHHHHHHH 77

RESULT 8
 B49122
 homeobox protein Mox-2 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 19-Dec-1993 #sequence_revision 17-Nov-1995 #text_change 24-Sep-1999
 C:Accession: B49122; S41779
 R:Candia, A.F.; Hu, J.; Crosby, J.; Lalle, P.A.; Nodén, D.; Nadeau, J.H.; Wright, C.V.
 Development 116, 1123-1136, 1992
 A>Title: Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed
 A:Reference number: A49122; MUID:93201999
 A:Accession: B49122
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-303 <CAN1>
 A:Cross-references: EMBL:Z16406; NID:957949; PIDN:CAAT8899.1; PID:957950
 A:Experimental source: C57BL
 A>Note: the complete translation is not shown
 R:Candia, A.F.; Kovalik, J.P.; Wright, C.V.E.
 Nucleic Acids Res. 21, 4982, 1993
 A>Title: Amino acid sequence of Mox-2 and comparison to its Xenopus and rat homologs.
 A:Reference number: S4169; MUID:94232829
 A:Accession: S41779
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-303 <CAN2>
 A:Cross-references: EMBL:Z16406; NID:957949; PIDN:CAAT8899.1; PID:957950
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:187-243/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0085;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 Db 67 GHHHHHHHHH 77

RESULT 9
 A48130
 growth arrest-specific homeobox protein Cvx - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
 C:Accession: A48130; S31976
 R:Gorski, D.H.; Lepage, D.F.; Patel, C.V.; Copeland, N.G.; Jenkins, N.A.; Walsh, K.
 Mol. Cell. Biol. 13, 3722-3733, 1993
 A>Title: Molecular cloning of a diverged homeobox gene that is rapidly down-regulated during growth arrest
 A:Reference number: A48130; MUID:93268321
 A:Accession: A48130
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-303 <GOR>

A:Cross-references: GB:Z17223; NID:957951; PIDN:CAAT8931.1; PID:957952
 A:Experimental source: Aorta
 A>Note: sequence extracted from NCBI backbone (NCBIN:132842; NCBI:132843)
 R:Gorski, D.H.; Lepage, D.F.; Patel, C.V.; Copeland, N.G.; Jenkins, N.A.; Walsh, K.
 Submitted to the EMBL Data Library, October 1992
 A>Description: Cvx: A diverged homeodomain gene which is rapidly down-regulated follo
 A:Reference number: S31976
 A:Accession: S31976
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-182, 'S', 184-303 <GO2>
 A:Cross-references: EMBL:Z17223
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:187-243/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0085;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 Db 67 GHHHHHHHHH 77

RESULT 10
 S39406
 homeotic protein otx1 - human
 C:Species: Homo sapiens (man)
 C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 15-Oct-1999
 C:Accession: S39406
 R:Strome, A.; Acampora, D.; Mallamaci, A.; Stornaiuolo, A.; d'Apice, M.R.; Nigro, V.
 EMBO J. 12, 2735-2747, 1993
 A>Title: A vertebrate gene related to orthodenticle contains a homeodomain of the bic
 A:Reference number: S33345; MUID:93327763
 A:Accession: S39406
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-354 <SIM>
 C:Genetics:
 A:Gene: otx1
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:39-95/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 354;
 Best Local Similarity 100.0%; Pred. No. 0.0098;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 Db 291 GHHHHHHHHH 301

RESULT 11
 S35345
 otx1 protein - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
 C:Accession: S35345
 R:Strome, A.; Acampora, D.; Mallamaci, A.; Stornaiuolo, A.; d'Apice, M.R.; Nigro, V.
 EMBO J. 12, 2735-2747, 1993
 A>Title: A vertebrate gene related to orthodenticle contains a homeodomain of the bic
 A:Reference number: S35345; MUID:93327763
 A:Accession: S35345
 A:Molecule type: DNA
 A:Residues: 1-355 <SIM>
 C:Genetics:
 A:Gene: otx1
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:39-95/Domain: homeobox homology <HOX>

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 355;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHNNHHNNH 12
|||||
DB 291 GHHNNHHNNH 301

RESULT 12

homeodomain protein ctx1 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999
C:Accession: I56547
R:Frantz, G.D.; Weimann, J.M.; Levin, M.E.; McConnell, S.K.
J. Neurosci. 14, 5725-5740, 1994
A:Title: Ctx1 and Ctx2 define layers and regions in developing cerebral cortex and cereb
A:Reference number: I56547; MUID:95016961
A:Accession: I56547
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: GB:I32602; NID:9535739; PIDN:AAA53557.1; PID:9535740
C:Genetics:
A:Gene: CTX1
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:39-95/Domain: homeobox homology <HOX>

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 355;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHNNHHNNH 12
|||||
DB 291 GHHNNHHNNH 301

RESULT 13

transcription repressor protein YY1 - human
N:Alternate names: transcription repressor/activator protein NF-E1
C:Species: Homo sapiens (man)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 08-Oct-1999
C:Accession: A40350; S78494; S33712; A56419
R:Shi, Y.; Seto, E.; Chang, L.S.; Shenk, T.
Cell 67, 377-388, 1991
A:Title: Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and re
A:Reference number: A40350; MUID:92005716
A:Accession: A40350
A:Molecule type: mRNA
A:Residues: 1-414 <SHI>
A:Cross-references: GB:M77698; NID:9186767; PIDN:AAA59467.1; PID:9186768
A:Experimental source: HeLa cells
A>Note: The authors translated the codon CGC for residue 371 as Lys, CGA for residue 375
R:Whitson, R.H.; Huang, T.; Dang, J.; Itakura, K.
Submitted to the EMBL Data Library, July 1992
A:Description: Observed and predicted DNA binding of a zinc finger protein which recogni
A:Reference number: S78494
A:Accession: S78494
A:Molecule type: mRNA
A:Residues: 1-195, G', 197-414 <WHI>
A:Cross-references: EMBL:Z14077; NID:938010; PIDN:CAA78455.1; PID:938011
R:Park, K.; Atchison, M.L.
Proc. Natl. Acad. Sci. U.S.A. 88, 9804-9808, 1991
A:Title: Isolation of a candidate repressor/activator, NF-E1 (YY-1), delta, that binds t
A:Reference number: A56419; MUID:92052179
A:Accession: S33712

A:Molecule type: mRNA
A:Residues: 1-64, R', 66-195, G', 197-414 <PAR>
A:Cross-references: GB:M76541; NID:9189173; PIDN:AAA59926.1; PID:9189174
C:Keywords: DNA binding; transcription regulation; zinc finger
F:298-330/Region: zinc finger CCH motif
F:327-347/Region: zinc finger CCH motif
F:355-377/Region: zinc finger CCH motif
F:385-407/Region: zinc finger CCH motif

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 414;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHNNHHNNH 12
|||||
DB 69 GHHNNHHNNH 79

RESULT 14

delta/YY1/NF-E1/UCBP transcription factor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: A48273; A42055
R:Saitany, G.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 90, 5559-5563, 1993
A:Title: Characterization of the mouse gene that encodes the delta/YY1/NF-E1/UCBP tr
A:Reference number: A48273; MUID:93296177
A:Accession: A48273
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-414 <RES>
A:Cross-references: GB:I13968; NID:9293847; PIDN:AAA40477.1; PID:9293849
R:Flanagan, J.R.; Becker, K.G.; Emilst, D.L.; Gleason, S.L.; Driggers, P.R.; Levi, B.
Mol. Cell. Biol. 12, 38-44, 1992
A:Title: Cloning of a negative transcription factor that binds to the upstream conser
A:Reference number: A42055; MUID:92107191
A:Accession: A42055
A:Molecule type: mRNA
A:Residues: 1-414 <FLA>
A:Cross-references: GB:M73963; NID:9202270; PIDN:AAA40522.1; PID:9202271
A>Note: sequence extracted from NCBI backbone (NCBIN:74641, NCBIPI:74642)
C:Genetics:
A:Introns: 227/1, 281/2, 301/3, 354/3

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 414;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHNNHHNNH 12
|||||
DB 70 GHHNNHHNNH 80

RESULT 15

transcription factor delta - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 05-Nov-1999
C:Accession: A56418
R:Hartharan, N.; Kelley, D.E.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 88, 9799-9803, 1991
A:Title: delta, a transcription factor that binds to downstream elements in several p
A:Reference number: A56418; MUID:92052178
A:Accession: A56418
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-414 <HAR>
A:Cross-references: GB:M74590; NID:9192940; PIDN:AAA37521.1; PID:9192941
C:Keywords: transcription factor; zinc finger

Query Match 2.18; Score 11; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GHHHHHHHHH 12
Db 70 GHHHHHHHHH 80

Search completed: August 22, 2000, 17:43:27
Job time: 1131 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:42:01 ; Search time 43.07 Seconds

(without alignments)
374.863 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGNHHHHHHSSGHIDDD.....RFIQKDYDVMQRLDDTSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 85661 seqs, 30989116 residues

Word size : 0

Total number of hits satisfying chosen parameters: 85661

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	36.7	326	1	LCRV_YERPE
2	170	32.6	170	1	CARL_YERPE
3	82	15.7	326	1	LCRV_YERPS
4	11	2.1	303	1	MOX2_HUMAN
5	11	2.1	303	1	MOX2_MOUSE
6	11	2.1	303	1	MOX2_MOUSE
7	11	2.1	354	1	OTX1_HUMAN
8	11	2.1	355	1	OTX1_MOUSE
9	11	2.1	355	1	OTX1_MOUSE
10	11	2.1	414	1	TYR1_HUMAN
11	11	2.1	414	1	TYR1_MOUSE
12	11	2.1	466	1	ZIC3_MOUSE
13	11	2.1	467	1	ZIC3_HUMAN
14	11	2.1	611	1	SNF1_CANGA
15	11	2.1	619	1	SNF1_CANTR
16	10	1.9	298	1	MOX2_XENLA
17	10	1.9	331	1	HXAL_MOUSE
18	10	1.9	335	1	HXAL_MOUSE
19	10	1.9	351	1	HRPX_PLAHO
20	10	1.9	375	1	CAR2_DICDI
21	10	1.9	387	1	QIN_AVIS3
22	10	1.9	390	1	PC_DROME
23	10	1.9	410	1	BR3B_HUMAN
24	10	1.9	411	1	BR3B_MOUSE
25	10	1.9	430	1	NOR2_RAT
26	10	1.9	449	1	GATE_HUMAN
27	10	1.9	451	1	BFI_CHICK
28	10	1.9	469	1	BFI_HUMAN
29	10	1.9	477	1	BFI_HUMAN
30	10	1.9	480	1	BFI_MOUSE
31	10	1.9	481	1	BFI_MOUSE
32	10	1.9	620	1	SNF1_CANAL
33	10	1.9	623	1	PNT1_DROME

34	10	1.9	626	1	NR43_HUMAN	092570 homo sapien
35	10	1.9	628	1	NR43_RAT	P51179 rattus norv
36	10	1.9	633	1	SNF1_YEAST	P06782 saccharomyc
37	10	1.9	657	1	KNOB_PLAEN	P06719 plasmodium
38	10	1.9	665	1	LAMA_XENLA	P11048 xenopus lae
39	10	1.9	717	1	UBPB_YEAST	P36026 saccharomyc
40	10	1.9	763	1	DIRK_HUMAN	Q13627 homo sapien
41	10	1.9	763	1	DIRK_MOUSE	O61214 mus musculu
42	10	1.9	763	1	YRK_RAT	O63470 rattus norv
43	10	1.9	950	1	URB1_USTMA	P40349 ustilago ma
44	10	1.9	1312	1	PIP1_DROME	P23455 drosophila
45	10	1.9	1356	1	KOM2_YEAST	P51862 saccharomyc
46	10	1.9	2164	1	CCAA_MOUSE	P97445 mus musculu
47	10	1.9	2212	1	CCAA_RAT	P54282 rattus norv
48	10	1.9	143	1	CCAA_HUMAN	O00555 homo sapien
49	10	1.9	143	1	Y970_METJA	O58380 methanococc
50	9	1.7	332	1	DLX2_MOUSE	P40764 mus musculu
51	9	1.7	333	1	HNK1_RAT	O08656 rattus norv
52	9	1.7	359	1	HNK1_MAIZE	P24345 zea mays (m
53	9	1.7	421	1	BR3A_MOUSE	P17208 mus musculu
54	9	1.7	423	1	BR3A_HUMAN	O01851 homo sapien
55	9	1.7	428	1	FKR4_MOUSE	O64733 mus musculu
56	9	1.7	469	1	A2AC_DIDMA	P35405 didelphis m
57	9	1.7	473	1	KNOB_PLAFA	P13817 plasmodium
58	9	1.7	568	1	DISC_DROME	P23792 drosophila
59	9	1.7	597	1	STAR_DROME	P42519 drosophila
60	9	1.7	634	1	KNOB_PLARG	P09346 plasmodium
61	9	1.7	2044	1	CCAH_HUMAN	O95180 homo sapien
62	9	1.7	2254	1	CCAG_RAT	O54898 rattus norv
63	9	1.7	2424	1	CCAA_RABIT	P27884 oryctolagus
64	9	1.5	59	1	HPN_HELPY	O48251 helicobacte
65	8	1.5	323	1	MAF1_MOUSE	P54841 mus musculu
66	8	1.5	323	1	MAF1_RAT	P54842 rattus norv
67	8	1.5	323	1	OTX1_BRARE	O91994 brachydanio
68	8	1.5	351	1	CAN2_CAEEL	O18879 caenorhabdi
69	8	1.5	409	1	BR11_BRARE	O90436 brachydanio
70	8	1.5	437	1	SP12_BRARE	P56224 brachydanio
71	8	1.5	486	1	MEC2_HUMAN	P51608 homo sapien
72	8	1.5	495	1	BRN1_MOUSE	P31361 mus musculu
73	8	1.5	497	1	BRN1_RAT	O63262 rattus norv
74	8	1.5	500	1	BRN1_HUMAN	P20264 homo sapien
75	8	1.5	508	1	CROC_DROME	P33027 drosophila
76	8	1.5	530	1	ZIC2_MOUSE	O63520 mus musculu
77	8	1.5	609	1	OPA_DROME	P39768 drosophila
78	8	1.5	684	1	EPB4_HCMVA	P17151 human cytom
79	8	1.5	972	1	TOP1_DROME	P30189 drosophila
80	8	1.5	1133	1	ATX9_FETTH	O95050 tetrahymena
81	8	1.5	1597	1	SOL_DROME	P27398 drosophila
82	8	1.3	16	1	LPH1_ECOLI	P03058 escherichia
83	7	1.3	91	1	S10B_BOVIN	P02638 bos taurus
84	7	1.3	91	1	S10B_HUMAN	P04271 homo sapien
85	7	1.3	91	1	S10B_MOUSE	P05114 mus musculu
86	7	1.3	91	1	S10B_RAT	P04631 rattus norv
87	7	1.3	97	1	VE7_HPV23	P50781 human papil
88	7	1.3	100	1	VE7_HPV38	O80908 human papil
89	7	1.3	104	1	PTLA_STRW	P26426 streptococc
90	7	1.3	105	1	PTLA_LACLA	P23532 lactococcus
91	7	1.3	111	1	CB21_CARPL	P38582 carnobacter
92	7	1.3	114	1	ASR2_LYCES	P37219 lycopersico
93	7	1.3	127	1	RL21_SYNP7	O94385 synchococc
94	7	1.3	143	1	MOCR_RHISH	P55363 rhizobium s
95	7	1.3	143	1	YLAB_ENTHR	O34913 entrococcu
96	7	1.3	147	1	FER5_METJA	O57699 methanococc
97	7	1.3	153	1	ENRP_BPP21	P27358 bacteriophag
98	7	1.3	162	1	RN_BACIN	P00649 bacillus in
99	7	1.3	162	1	RN_BACPU	P48068 bacillus pu
100	7	1.3	182	1	DYR_DROME	P71719 drosophila
101	7	1.3	193	1	YDGM_HAEIN	P71396 haemophilus
102	7	1.3	216	1	UREY_RHIME	P47880 rhizobium m
103	7	1.3	252	1	COR2_EMENI	P13588 emeticella
104	7	1.3	253	1	KDSB_HAEIN	P44490 haemophilus
105	7	1.3	297	1	CTAA_BACET	O04443 bacillus fi
106	7	1.3	302	1	HYPB_BRAJA	O45257 bradyrhizob

107	1.3	321	1	CANC_T0BAC	P27111	nitotiana t	180	1.2	65	1	TRBK_RHISN	P55401	rhizobium s
108	1.3	328	1	DLX2_HUMAN	Q07687	homo sapien	181	1.2	70	1	ATP1_BOVIN	Q00361	bos taurus
109	1.3	328	1	OPY1_YEAST	P38221	saccharomyc	182	1.2	78	1	KCR9_PIG	Q29577	sus scrofa
110	1.3	329	1	CAHX_FLAPR	P46281	flaviera pr	183	1.2	80	1	COXK_MOUSE	P56992	mus musculus
111	1.3	330	1	CAH1_FLALI	P46512	flaviera il	184	1.2	80	1	PSY1_SYNP2	P19165	synecchoc
112	1.3	330	1	CAHX_FLABR	P46511	flaviera br	185	1.2	82	1	YB1M_HERRA	P23777	hepeticosiph
113	1.3	331	1	CAHX_FLABI	P46510	flaviera bl	186	1.2	82	1	YB1M_HERRA	P23777	hepeticosiph
114	1.3	332	1	G3P3_ECOLI	P33898	escherichia	187	1.2	85	1	YOR4_BPSPP	P25281	hepeticosiph
115	1.3	336	1	GUNA_RUMFL	P16139	ruminococcus	188	1.2	89	1	TUBE_MYCTU	Q38440	bacterioph
116	1.3	348	1	YHHT_HABIN	P46466	haemophilus	189	1.2	92	1	RPC1_BPP22	P02444	bacterioph
117	1.3	350	1	HMRO_DROME	P10181	drosophila	190	1.2	92	1	Y766_RICPR	P03041	bacterioph
118	1.3	350	1	Y098_SYNY3	Q05886	synecchocyst	191	1.2	93	1	VE7_HPV50	Q92635	rickettsia
119	1.3	359	1	Y0R3_CAREL	Q09580	caenorhabdit	192	1.2	95	1	NOLE_RHILP	Q80928	human papil
120	1.3	362	1	VM25_BORHE	P32779	caenorhabdit	193	1.2	99	1	BZMG_RABIT	P23717	rhizobium l
121	1.3	364	1	VM21_BORHE	P21875	borrelia he	194	1.2	100	1	VE7_HPV22	P01885	oryctolagus
122	1.3	374	1	YDVB_SCHPO	O14229	schistosach	195	1.2	101	1	VG7_BPP03	P50780	human papil
123	1.3	374	1	YMP4_CAREL	Q10948	caenorhabdit	196	1.2	103	1	PTLA_STRAU	Q37887	bacterioph
124	1.3	376	1	T2A4_HUMAN	P52655	homo sapien	197	1.2	106	1	Y115_ADE07	P02909	staphylococ
125	1.3	380	1	CAH2_CHLRE	P24238	chlamydomon	198	1.2	107	1	SS1_STRGI	P03288	human adeno
126	1.3	385	1	SELP_RAT	P25236	rattus norv	199	1.2	111	1	YCX1_CHLRE	P28392	streptomyce
127	1.3	392	1	PURT_PASHA	P46927	p. probable	200	1.2	112	1	VPX_HVZD1	P05723	chlamydomon
128	1.3	398	1	TBX1_HUMAN	Q43435	homo sapien	201	1.2	112	1	VPX_HVZG1	P17760	human immun
129	1.3	401	1	PRRB_ECOLI	P17222	escherichia	202	1.2	112	1	Y13K_SSV1	P18045	human immun
130	1.3	408	1	FRF2_HUMAN	Q12947	homo sapien	203	1.2	113	1	PSAK_CHLRE	P20920	sulfolobus
131	1.3	410	1	T1S1_ECOLI	P10465	escherichia	204	1.2	113	1	VC21_VACCC	P14225	chlamydomon
132	1.3	425	1	POU1_BRARE	P31366	brachydanio	205	1.2	113	1	VPX_HV2BE	P21105	vaccinia vi
133	1.3	441	1	GAT6_RAT	P46153	rattus norv	206	1.2	121	1	KG20_PYRAB	P18099	human immun
134	1.3	443	1	ZP23_BRARE	P79745	brachydanio	207	1.2	121	1	YD60_PYRAB	P56813	pyrococcus
135	1.3	444	1	GAT6_MOUSE	Q61169	mus musculus	208	1.2	121	1	LCA_MACRG	P00556	escherichia
136	1.3	485	1	CG2A_CAREL	P34638	caenorhabdit	209	1.2	122	1	GSTB_BACSU	P07458	macropus ru
137	1.3	480	1	HMS1_YEAST	P54725	saccharomyc	210	1.2	125	1	YOR3_YEREN	P26307	bacillus su
138	1.3	492	1	MEC2_RAT	Q00566	rattus norv	211	1.2	127	1	YQUB_ECOLI	P28836	yeastinia su
139	1.3	515	1	PPZ_SCHPO	P78968	schizosacch	212	1.2	130	1	SZ05_RAT	P97885	escherichia
140	1.3	518	1	YLM6_CAREL	P34409	caenorhabdit	213	1.2	131	1	VE98_AQUAE	Q48405	bacterioph
141	1.3	519	1	SAZD_HUMAN	Q12788	homo sapien	214	1.2	131	1	YEP51_BPSPI	Q67469	aquifex deo
142	1.3	532	1	VIL1_HPV03	P36731	human papil	215	1.2	133	1	EXD1_HELPY	O25754	helicobacte
143	1.3	535	1	DPPA_ECOLI	P23847	escherichia	216	1.2	135	1	ACP_BRACM	P07088	diassilia ca
144	1.3	570	1	NU5M_PARTE	P13584	paramecium	217	1.2	135	1	V47_BPP7	P03786	bacterioph
145	1.3	572	1	PT1_STAUV	P51183	staphylococ	218	1.2	140	1	LCA_MACEU	Q06555	macropus eu
146	1.3	579	1	G160_HUMAN	Q08378	homo sapien	219	1.2	140	1	PRO3_WHEAT	P49324	tritlicum ae
147	1.3	587	1	HE_PARI1	P22757	paracentrot	220	1.2	140	1	VHM3_CAPVK	P16718	capripoxvir
148	1.3	604	1	Y1J3_YEAST	P47030	saccharomyc	221	1.2	141	1	PRO2_WHEAT	P49333	tritlicum ae
149	1.3	615	1	AIPI_YEAST	P46860	saccharomyc	222	1.2	142	1	RR8_ASTRO	P24353	astasia ion
150	1.3	617	1	YA42_HABIN	Q57175	haemophilus	223	1.2	144	1	MARR_SALTY	Q56069	salmonella
151	1.3	627	1	CIAT_CAREL	P32756	caenorhabdit	224	1.2	144	1	VGR_BPP1	P22446	bacterioph
152	1.3	682	1	SNK_MOUSE	P53351	mus musculus	225	1.2	144	1	YRA1_BACSU	Q07909	bacillus su
153	1.3	746	1	TDCE_ECOLI	P42632	escherichia	226	1.2	148	1	RPR8_CAREL	O19826	caenorhabdit
154	1.3	785	1	SECA_MYCLE	Q32932	mycobacteri	227	1.2	148	1	SSB_HUMAN	Q04837	homo sapien
155	1.3	785	1	YGS1_YEAST	P50089	saccharomyc	228	1.2	151	1	SSB_RAT	P28042	rattus norv
156	1.3	808	1	SEAZ_MYCTU	Q50612	mycobacteri	229	1.2	155	1	VPR_BPP2	P36933	bacterioph
157	1.3	817	1	HUNB_MUSDO	Q01778	musca domes	230	1.2	158	1	GREA_ECOLI	P21446	bacterioph
158	1.3	836	1	PDC2_CANAL	O60035	candida alb	231	1.2	159	1	SMRA_TREHY	Q54145	treponema h
159	1.3	886	1	MCM2_XENLA	P55861	xenopus lae	232	1.2	161	1	UREX_PROMI	P17900	proteus mir
160	1.3	887	1	MCM2_DROME	P49735	drosophila	233	1.2	162	1	PHAC_SYNP6	P11390	synecchococ
161	1.3	895	1	MCM2_HUMAN	P49736	homo sapien	234	1.2	162	1	RS16_MYCTU	Q10795	mycobacteri
162	1.3	904	1	MCM2_MOUSE	P97310	mus musculus	235	1.2	163	1	ATPX_GUTHH	Q78478	halobacteria
163	1.3	906	1	CB1_HUMAN	P22661	homo sapien	236	1.2	163	1	GVH2_HALSA	P33361	halobacteria
164	1.3	944	1	MMU3_MYCTU	O53657	mycobacteri	237	1.2	164	1	LSPA_ECOLI	P00804	escherichia
165	1.3	946	1	YBTE_YEAST	P38282	saccharomyc	238	1.2	165	1	LSPA_EYTAE	P13314	enterobacte
166	1.3	946	1	YBTE_YEAST	P40559	saccharomyc	239	1.2	166	1	RNH2_LACLC	Q30415	lactococcus
167	1.3	946	1	YBTE_YEAST	Q00808	podospora a	240	1.2	166	1	YN21_CAREL	P45561	caenorhabdit
168	1.3	1356	1	HET1_PODAN	O00808	podospora a	241	1.2	169	1	BENB_ACICA	P07770	acinetobact
169	1.3	1705	1	CYAA_BORBE	O57506	bordeletia	242	1.2	170	1	GBP_XENLA	Q93343	xenopus lae
170	1.3	1706	1	CYAA_BORBE	P15318	bordeletia	243	1.2	170	1	LSPA_PEEFL	P17942	pseudomonas
171	1.3	1912	1	PTPD_HUMAN	P23468	homo sapien	244	1.2	172	1	APT_STAUV	O32418	staphylococ
172	1.3	2326	1	CCAB_DISOM	P56698	drosophila	245	1.2	173	1	PASE_ECOLI	P08407	escherichia
173	1.3	2483	1	PCX_DROME	P18450	drosophila	246	1.2	173	1	PRSE_ECOLI	P42186	escherichia
174	1.2	16	1	TRYP_FELCA	P81071	felis silve	247	1.2	175	1	RP18_MOUSE	P47939	mus musculus
175	1.2	31	1	PSBT_CHLRE	P57256	chlamydomon	248	1.2	175	1	IPYR_THEAC	P37981	thermoplasm
176	1.2	31	1	PSBT_CHLRE	P56337	chlorella v	249	1.2	179	1	RL5_BACSU	P12877	macillus su
177	1.2	35	1	PSBT_MARPO	P12182	marichantia	250	1.2	179	1	CYPA_MYCTU	P71378	mycobacteri
178	1.2	54	1	IYOVO_CASCA	P05559	casuarina c	251	1.2	182	1	K2C3_BOVIN	P04261	bos taurus
179	1.2	54	1	IYOVO_DRONO	P05560	diromalus no	252	1.2	182	1			

253	6	1.2	182	1	K2C4_BOVIN	P04260	bos taurus
254	6	1.2	184	1	P12_LUPO	P16148	lupinus pol
255	6	1.2	184	1	TRX1_MOUSE	P70323	mus musculus
256	6	1.2	185	1	NIR_LEPMC	P45004	leptosphaer
257	6	1.2	192	1	PABA_STRLI	P27637	streptomyces
258	6	1.2	192	1	Y810_BORBU	P70837	botryella bu
259	6	1.2	193	1	Y983_HAETN	P43907	haemophilus
260	6	1.2	194	1	RS7_ARCFU	O28386	archaeoglob
261	6	1.2	195	1	INT_CAPHI	P28171	capra hircu
262	6	1.2	197	1	YGVV_ECOLI	P50061	escherichia
263	6	1.2	198	1	APPF_MYCGA	P33256	mycoplasma
264	6	1.2	198	1	FRS6_RHORA	P43378	rhodococcus
265	6	1.2	198	1	YCXI_PORPU	P51354	porphyra pu
266	6	1.2	199	1	PA10_YEAST	P48363	saccharomyces
267	6	1.2	201	1	YAMC_SCHPO	O10186	schizosacch
268	6	1.2	204	1	YAS7_ACTAC	O57728	actinobacill
269	6	1.2	205	1	RNH2_ARCFU	O28634	archaeoglob
270	6	1.2	208	1	Y4IN_RHISN	P53497	rhizobium s
271	6	1.2	209	1	GC_NETGO	O87406	neisseria g
272	6	1.2	210	1	CD8B_PONPY	P30434	pongo pygma
273	6	1.2	210	1	YG21_YEAST	P53251	saccharomyces
274	6	1.2	212	1	PECF_MASIA	P29730	mastigoclad
275	6	1.2	212	1	KERC_ARATH	O94W17	arabidopsis
276	6	1.2	213	1	HIS1_BACSU	O34520	bacillus su
277	6	1.2	213	1	HPRT_MUSSP	O64531	mus spretus
278	6	1.2	217	1	HPRT_CRIGR	P00494	cricketulus
279	6	1.2	217	1	HPRT_MERUN	P47959	meriones un
280	6	1.2	217	1	HPRT_MOUSE	P00493	mus musculus
281	6	1.2	218	1	HPRT_RAT	P27605	rattus norv
282	6	1.2	219	1	THIE_HELPY	O25514	helicobacte
283	6	1.2	220	1	DEOC_MYCPI	P47722	mycoplasma
284	6	1.2	220	1	YGCA_VIBSS	P55136	vibrio sp.
285	6	1.2	221	1	CART_MOUSE	O08331	mus musculus
286	6	1.2	223	1	YEL2_EBV	P03203	epstein-bar
287	6	1.2	224	1	KTHX_CHICK	P04047	gallus galli
288	6	1.2	226	1	Y703_METJA	O58114	methanococ
289	6	1.2	230	1	FLAF_METYO	O06641	methanococ
290	6	1.2	231	1	PSAF_SPIOL	P12355	spinacia ol
291	6	1.2	233	1	H1I_GLYXA	P40263	glyptotendi
292	6	1.2	233	1	H1I_GLYXA	P40266	glyptotendi
293	6	1.2	234	1	RK1_GUTTH	O78413	guillardia
294	6	1.2	234	1	Y6IB_ECOLI	P24195	escherichia
295	6	1.2	239	1	TRPA_THEMA	P50908	thermotoga
296	6	1.2	240	1	YB20_PLAFA	O96178	plasmidium
297	6	1.2	243	1	IP72_AGR7	P06524	agrobacteri
298	6	1.2	243	1	IPR_AGR7	P14011	agrobacteri
299	6	1.2	243	1	YHFR_ECOLI	P45544	escherichia
300	6	1.2	244	1	TONB_VIBCH	O50442	vibrio chol
301	6	1.2	246	1	TRY1_CANPA	P06871	canis famli
302	6	1.2	247	1	ATP6_ACACA	O37385	acanthamoeb
303	6	1.2	248	1	YIPA_YEAST	P53039	saccharomyces
304	6	1.2	249	1	YCIT_ECOLI	P70034	escherichia
305	6	1.2	250	1	Y035_METJA	O60350	methanococ
306	6	1.2	252	1	DEOR_ECOLI	P06217	escherichia
307	6	1.2	252	1	ONCM_HUMAN	P13725	homo sapien
308	6	1.2	252	1	TRPC_THEMA	O56319	thermotoga
309	6	1.2	252	1	YR2_CAEEL	O10006	caenorhabdi
310	6	1.2	253	1	UPPS_ECOLI	P4434	arabidopsis
311	6	1.2	253	1	SSRA_ARATH	O47675	escherichia
312	6	1.2	253	1	YTBQ_BACSU	P53560	bacillus su
313	6	1.2	253	1	YX04_CAEEL	O11111	caenorhabdi
314	6	1.2	255	1	MYB3_MAZE	P20025	zee mays (m
315	6	1.2	256	1	YKHI_YEAST	P36086	saccharomyces
316	6	1.2	257	1	ETXA_STAUD	P13163	staphylococ
317	6	1.2	258	1	STV_STRCO	O06851	streptomyces
318	6	1.2	258	1	Y098_MYCTU	O10859	mycobacteri
319	6	1.2	261	1	VP31_FRG3V	P18178	frog virus
320	6	1.2	264	1	FLGG_BACSU	P23446	bacillus su
321	6	1.2	264	1	YJFZ_ECOLI	P39308	escherichia
322	6	1.2	267	1	Y617_SYNY3	O55707	synechocyst
323	6	1.2	268	1	CEBD_MOUSE	O00322	mus musculus
324	6	1.2	268	1	CEBD_RAT	O03484	rattus norv
325	6	1.2	268	1	FLHP_BACSU	P39753	bacillus su
326	6	1.2	269	1	CEBD_HUMAN	P49716	homo sapien
327	6	1.2	271	1	CART_HUMAN	P22676	homo sapien
328	6	1.2	271	1	CART_RAT	P47787	rattus norv
329	6	1.2	271	1	HMR_XENLA	P14883	xenopus lae
330	6	1.2	271	1	HXA3_MOUSE	P09631	mus musculus
331	6	1.2	271	1	VIUB_VIBCH	O56646	vibrio chol
332	6	1.2	271	1	YW34_MYCTU	O05879	mycobacteri
333	6	1.2	272	1	DKH1_STRVN	P16542	streptomyces
334	6	1.2	272	1	HXA9_HUMAN	P12659	homo sapien
335	6	1.2	273	1	FLJM_CAUCR	O52529	caulobacter
336	6	1.2	273	1	IACG_AGRD	P29874	agrobacteri
337	6	1.2	273	1	T2C2_HERAU	P25259	herpetosiph
338	6	1.2	273	1	Y133_METJA	O57537	methanococ
339	6	1.2	274	1	FLJO_CAUCR	O52531	caulobacter
340	6	1.2	274	1	T2B1_HERAU	P25257	herpetosiph
341	6	1.2	274	1	Y0HM_HERAU	P25260	herpetosiph
342	6	1.2	274	1	Y0HM_HERAU	P76425	escherichia
343	6	1.2	276	1	BACH_HALHP	O48315	halobacteri
344	6	1.2	280	1	YMEI_YEAST	O34433	saccharomyces
345	6	1.2	280	1	CEBE_HUMAN	O15744	homo sapien
346	6	1.2	281	1	CEBE_HUMAN	O58484	methanococ
347	6	1.2	282	1	RS2_MYCTU	O10796	mycobacteri
348	6	1.2	287	1	DCOP_MYXXA	P24230	mycobacteri
349	6	1.2	288	1	PIP_BACCO	P46541	bacillus co
350	6	1.2	288	1	YIGM_ECOLI	P77849	escherichia
351	6	1.2	288	1	ARY3_MOUSE	P50296	mus musculus
352	6	1.2	290	1	GDBB_WHEAT	P06659	mus musculus
353	6	1.2	291	1	NLA_DROME	O94W17	arabidopsis
354	6	1.2	292	1	YQFU_BACSU	P24478	bacillus su
355	6	1.2	293	1	RS2_MYCPN	P75560	mycoplasma
356	6	1.2	294	1	MED6_YEAST	P87872	saccharomyces
357	6	1.2	295	1	PERX_BRARA	P00434	brassica ra
358	6	1.2	296	1	YD28_METJA	O58774	methanococ
359	6	1.2	298	1	DAPA_RHIME	O59762	rhizobium m
360	6	1.2	300	1	APBA_PYRHO	O50098	pyrococcus
361	6	1.2	301	1	P061_HUMAN	O14863	homo sapien
362	6	1.2	301	1	P061_MOUSE	O07916	mus musculus
363	6	1.2	301	1	P061_MOUSE	P51232	rattus norv
364	6	1.2	302	1	GDBX_WHEAT	O50559	methanobact
365	6	1.2	303	1	KIME_MERTH	P14287	sus scrofa
366	6	1.2	303	1	OSTP_PIG	O12336	saccharomyces
367	6	1.2	303	1	PMG3_YEAST	P24280	saccharomyces
368	6	1.2	303	1	SC14_YEAST	O44104	dirosophila
369	6	1.2	304	1	G3P2_DROPS	O44105	dirosophila
370	6	1.2	304	1	G3P2_DROPS	P31316	mus musculus
371	6	1.2	305	1	GS2_MOUSE	P53389	rhizobium s
372	6	1.2	305	1	Y4CG_RHISN	P45084	haemophilus
373	6	1.2	306	1	OPPB_HAETN	P25666	pseudomonas
374	6	1.2	306	1	SDSB_PSES9	P57578	escherichia
375	6	1.2	306	1	YKIC_ECOLI	P61615	saccharomyces
376	6	1.2	307	1	VE2_HPV5	P67578	human papil
377	6	1.2	310	1	PMG2_YEAST	O12008	saccharomyces
378	6	1.2	311	1	TRAD_HUMAN	O15628	homo sapien
379	6	1.2	312	1	ETFA_BRARA	P33733	bradyrhizob
380	6	1.2	314	1	YNO3_YEAST	P33845	saccharomyces
381	6	1.2	314	1	PER1_ARAHY	P22155	atachias hyp
382	6	1.2	316	1	THFR_SYNP7	P27477	synechococ
383	6	1.2	320	1	ARGC_ANASP	P44894	anabaena sp
384	6	1.2	322	1	ANXS_CYNPY	P70075	cynops pyr
385	6	1.2	323	1	ARGI_EMDNI	O12611	emeritella
386	6	1.2	323	1	CCG2_HUMAN	O56698	homo sapien
387	6	1.2	323	1	CCG2_MOUSE	O88602	mus musculus
388	6	1.2	323	1	ISPB_ECOLI	P19641	escherichia
389	6	1.2	323	1	Y0XA_BACSU	P39840	bacillus su
390	6	1.2	325	1	T2ER_YEAST	P61455	saccharomyces
391	6	1.2	328	1	Y36K_HALSP	P14331	halobacteri
392	6	1.2	328	1	BIOB_BACSH	P19260	bacillus sp
393	6	1.2	332	1	G1P1_DROME	P07466	dirosophila
394	6	1.2	332	1	G3P2_DROME	P07487	dirosophila
395	6	1.2	332	1	G3P_DROME	O01597	dirosophila
396	6	1.2	332	1	YG2H_YEAST	P53250	saccharomyces
397	6	1.2	332	1			
398	6	1.2	332	1			

399	6	1.2	333	1	MP92_POAPR	P2285	poa pletens	472	6	1.2	377	1	GB13_HUMAN	Q1344	homo sapien
400	6	1.2	334	1	G3P1_ANAVA	P3416	anabaena va	473	6	1.2	377	1	YPER_BACSU	P54169	baicillus su
401	6	1.2	334	1	VAS1_VACCC	P21069	vaccinia vl	474	6	1.2	378	1	CYB_ELEMA	O4785	elephas max
402	6	1.2	334	1	VAS1_VACCV	Q01119	vaccinia vl	475	6	1.2	378	1	NMT3_CANAL	P87207	candida alb
403	6	1.2	334	1	VAS1_VARV	P33858	variolae vir	476	6	1.2	378	1	YVNB_YEAST	P55923	saccharomyc
404	6	1.2	336	1	Y883_METJA	O5893	methanococ	477	6	1.2	379	1	CYB_LOXAP	P24958	toxodonta a
405	6	1.2	338	1	DCUP_AOUAE	O6667	aquifex aeo	478	6	1.2	379	1	NHBI_SOYBN	P24958	toxodonta a
406	6	1.2	338	1	SIFA_AOUAE	O6787	aquifex aeo	479	6	1.2	379	1	YAXO_RHISN	P4608	glycine max
407	6	1.2	338	1	TRSS_ECOLI	P03837	escherichia	480	6	1.2	381	1	CHYM_BOVIN	P55707	rhizobium s
408	6	1.2	339	1	SPSG_BACSU	P39627	baicillus su	481	6	1.2	381	1	CHYM_SHEEP	P00794	bos taurus
409	6	1.2	339	1	SRPA_SYNP7	O55025	synchococ	482	6	1.2	383	1	PO13_NASVT	O03271	ovis aries
410	6	1.2	340	1	YLS1_YEAST	Q12524	saccharomyc	483	6	1.2	385	1	DHAB_HUMAN	P48448	homo sapien
411	6	1.2	341	1	ETRA_SCHPO	P78790	schizosacch	484	6	1.2	390	1	CBT_MLVCN	P23092	cas-n8-1 mu
412	6	1.2	342	1	RECA_EKACA	P26344	erwinia car	485	6	1.2	390	1	RPSD_ANASP	P26683	anabaena sp
413	6	1.2	344	1	MURG_AOUAE	O67238	aquifex aeo	486	6	1.2	390	1	SAOX_BACBP	P40859	baicillus sp
414	6	1.2	344	1	TMRA_TREPH	P29721	treponema p	487	6	1.2	391	1	BCHP_RHOCA	P26172	rhodobacter
415	6	1.2	345	1	VP10_MTV	P13093	wound tumor	488	6	1.2	391	1	Y532_METJA	O57952	methanococ
416	6	1.2	347	1	F16P_YEAST	P09201	saccharomyc	489	6	1.2	392	1	CYSA_STRO	O58829	streptomyc
417	6	1.2	347	1	IDH_METJA	O58891	methanococ	490	6	1.2	392	1	ODD_DROME	P23803	strepomyce
418	6	1.2	348	1	RMBE_STRMU	P95780	streplococ	491	6	1.2	394	1	YJ51_YEAST	P47127	saccharomyc
419	6	1.2	349	1	ASG2_HAEIN	P43843	haemophilus	492	6	1.2	395	1	VPF1_BPP2	P22501	bacterioph
420	6	1.2	349	1	GPDA_DROAE	Q27556	escherichia	493	6	1.2	395	1	Y305_METJA	O57753	methanococ
421	6	1.2	349	1	YGB0_ECOLI	O57261	escherichia	494	6	1.2	396	1	REBX_SHIDY	O03583	shigella dy
422	6	1.2	352	1	GBA1_CRYPA	O00580	cryphonectr	495	6	1.2	396	1	SP1_HUMAN	O955X1	homo sapien
423	6	1.2	352	1	GPDA_DROVI	P07335	drosophila	496	6	1.2	397	1	SK18_YEAST	Q02793	saccharomyc
424	6	1.2	352	1	KAPC_DROME	P12370	drosophila	497	6	1.2	398	1	SME_SYNY3	P73345	synchocyst
425	6	1.2	352	1	Z185_MOUSE	Q62394	mus musculu	498	6	1.2	400	1	LEUK_HUMAN	P16150	homo sapien
426	6	1.2	354	1	PON2_MOUSE	O62086	mus musculu	499	6	1.2	400	1	MUAI_XENLA	P10667	xenopus lae
427	6	1.2	354	1	RECA_PROMI	P1406	proteus mir	500	6	1.2	400	1	TRAA_RHIME	P80011	rhizobium m
428	6	1.2	354	1	YMA3_MYCBO	O02279	mycobacteri	501	6	1.2	400	1	YAPO_RHISN	P55620	rhizobium s
429	6	1.2	355	1	AMIE_STRPN	P18765	streplococ	502	6	1.2	400	1	YCEI_BACSU	O34691	baicillus su
430	6	1.2	355	1	CMG2_SCHPO	P32434	schizosacch	503	6	1.2	401	1	P39_BRUAB	O06875	bruceella ab
431	6	1.2	355	1	LEB9_HUMAN	O00182	homo sapien	504	6	1.2	401	1	PGK_SYNY3	P74421	synchocyst
432	6	1.2	355	1	MTM2_METTF	P29568	methanobact	505	6	1.2	403	1	MHPT_ECOLI	P77589	escherichia
433	6	1.2	358	1	G3PG_TRYBB	P22112	trypanosoma	506	6	1.2	404	1	MEIK_CAEEL	P50303	caenorhabdi
434	6	1.2	358	1	RECA_XENBY	P96185	xenorhabdus	507	6	1.2	404	1	TY3H_CAEEL	P90986	caenorhabdi
435	6	1.2	358	1	SP11_MOUSE	O9WIK8	mus musculu	508	6	1.2	405	1	PRSB_DROME	O18413	drosophila
436	6	1.2	359	1	G3PG_TRYCR	P22113	trypanosoma	509	6	1.2	405	1	SDC3_CHICK	P26261	galinus galli
437	6	1.2	360	1	DCAM_SOLTU	Q04594	solanum tub	510	6	1.2	406	1	PRSB_HUMAN	P47210	homo sapien
438	6	1.2	360	1	MK14_HUMAN	Q16539	h. mitogen-a	511	6	1.2	406	1	PRSB_MOUSE	P52915	mus musculu
439	6	1.2	360	1	MK14_MOUSE	P47811	mus musculu	512	6	1.2	407	1	COAT_FHV	P18870	flock house
440	6	1.2	360	1	MK14_MOUSE	P70618	rattus norv	513	6	1.2	407	1	VG10_HSVSA	P24913	hepesvirus
441	6	1.2	360	1	UXRA_THEMA	O9WXS4	thermotoga	514	6	1.2	409	1	EFUO_ASTLO	P14634	astasia ion
442	6	1.2	361	1	VE2_HPV70	P50773	human papil	515	6	1.2	409	1	RHAG_HUMAN	Q02094	homo sapien
443	6	1.2	361	1	DCAM_NICSY	O80402	nicotiana s	516	6	1.2	410	1	AMPS_BACSU	P39762	baicillus su
444	6	1.2	362	1	DCAM_TOBAC	O04009	nicotiana t	517	6	1.2	410	1	HMH2_DROME	P10035	synchocyst
445	6	1.2	362	1	DCAM_DISTR	O96555	datura stra	518	6	1.2	410	1	PGIR_MAIZE	P26215	zea mays (m
446	6	1.2	363	1	Y593_CHLPN	O9ZTW1	chlamydia p	519	6	1.2	410	1	PGIS_MAIZE	P38338	zea mays (m
447	6	1.2	363	1	BAS5_ECOLI	P30844	escherichia	520	6	1.2	411	1	STRD_SOYBN	Q42807	glycine max
448	6	1.2	363	1	DCAM_EPIOL	P46255	sphingia ol	521	6	1.2	412	1	AAT_RICPR	O94556	ricicetia
449	6	1.2	363	1	MURG_BACSU	P37585	baicillus su	522	6	1.2	412	1	CAR2_CANPA	P33950	candida par
450	6	1.2	363	1	OMPF_SALTI	O56113	salmonella	523	6	1.2	413	1	FENR_SYNY3	O53118	synchocyst
451	6	1.2	363	1	OMPF_SALTY	P37432	salmonella	524	6	1.2	413	1	GAT1_RAT	P43429	rattus norv
452	6	1.2	366	1	SVY_SULSO	P95982	sulfolobus	525	6	1.2	416	1	CCA_HAEIN	P42269	hemophilus
453	6	1.2	369	1	HIP_HUMAN	P50802	homo sapien	526	6	1.2	417	1	IF_RAT	P17267	rattus norv
454	6	1.2	369	1	MAR2_RAT	P54844	rattus norv	527	6	1.2	417	1	KCRU_CHICK	P170079	galinus galli
455	6	1.2	369	1	SERC_HELIX	O25436	helicobacte	528	6	1.2	417	1	KCRU_HUMAN	P12532	homo sapien
456	6	1.2	369	1	TMAF_AVISA	P23091	avlan muscu	529	6	1.2	418	1	KCRU_MOUSE	P30275	mus musculu
457	6	1.2	369	1	VM07_BORHE	P21876	borrelia he	530	6	1.2	418	1	KCRU_MOUSE	P25809	rattus norv
458	6	1.2	369	1	YDHH_ECOLI	P77570	escherichia	531	6	1.2	419	1	Y092_HAEIN	O57493	hemophilus
459	6	1.2	370	1	EXOH_RHIME	P33692	rhizobium m	532	6	1.2	420	1	PEL_BACSU	P39116	baicillus su
460	6	1.2	370	1	MAR2_MOUSE	P54843	mus musculu	533	6	1.2	421	1	PELR_MEDSA	O40312	medicago sa
461	6	1.2	372	1	YBDK_ECOLI	P77213	escherichia	534	6	1.2	423	1	OPPD_MYCPN	P75552	mycoplasma
462	6	1.2	373	1	TGFI_CHICK	P09531	galus galli	535	6	1.2	423	1	SVS_COXBG	P33919	coxciella bu
463	6	1.2	374	1	FOF_TETFL	O91496	tetradon f	536	6	1.2	423	1	TIG_BACSU	P80698	baicillus su
464	6	1.2	374	1	GB15_HUMAN	P30679	homo sapien	537	6	1.2	424	1	ENO_CHLTR	O84591	chlamydia t
465	6	1.2	374	1	HAT1_YEAST	Q12241	saccharomyc	538	6	1.2	425	1	ENO_CHLTR	O84591	chlamydia t
466	6	1.2	374	1	YHNE_YEAST	P38797	saccharomyc	539	6	1.2	426	1	Y958_METJA	O27502	methanococ
467	6	1.2	375	1	Y785_METJA	O58195	methanococ	540	6	1.2	427	1	VP8_MTV	P17380	wound tumor
468	6	1.2	376	1	CISX_PYRFU	O53554	pyrococcus	541	6	1.2	427	1	FE1A_METJA	O57770	methanococ
469	6	1.2	376	1	FOS_FUGRU	P53450	fugu rubrip	542	6	1.2	428	1	ENO_CHLPN	O92766	chlamydia p
470	6	1.2	376	1	OMPC_SERMA	O54471	seirratia ma	543	6	1.2	428	1	ENO_NITEU	O85348	nltrosomona
471	6	1.2	377	1	CAH1_CHLRE	P20507	chlamydomon	544	6	1.2	428	1	ENO_NITEU	O85348	nltrosomona

545	6	1.2	435	1	CG65_YEAST	P30283	saccharomyc	618	6	1.2	488	1	PBB_ALCEA	P12625	alcaligenes
546	6	1.2	436	1	CP22_HORVU	P55748	hordium_vul	619	6	1.2	488	1	RBL1_CYACA	P37393	cyandium c
547	6	1.2	436	1	MTBR_BACSU	P06530	bacillifig su	620	6	1.2	488	1	RBL_ECTSI	P24313	ectocarpus
548	6	1.2	436	1	TDH1_CORGL	Q04513	corynebacte	621	6	1.2	488	1	RBL_GUTHI	P14957	guillierdia
549	6	1.2	438	1	MPK5_HUMAN	Q1163	homo sapien	622	6	1.2	488	1	RBL_OLITU	P14959	olisthodisc
550	6	1.2	442	1	TBB_TRXCR	P08562	trypanosoma	623	6	1.2	488	1	RBL_PLECA	Q08051	pleurochrys
551	6	1.2	443	1	FLI1_AQUAE	O67531	aquifex aeo	624	6	1.2	488	1	RBL_PORAE	O09119	porphyridi
552	6	1.2	443	1	YLM5_CAEEL	P34379	caenorhabdi	625	6	1.2	488	1	RBL_PORPU	P51226	porphyra pu
553	6	1.2	444	1	DHNA_HAEIN	P44856	haemophilus	626	6	1.2	488	1	RBL_PYLTI	P23651	pylataella i
554	6	1.2	444	1	PVDA_BURCE	O51940	burkholderi	627	6	1.2	488	1	VE2_HPV49	P36795	human papil
555	6	1.2	444	1	YNBL_YEAST	P53980	saccharomyc	628	6	1.2	489	1	FLIC_SALTU	P06179	salmonella
556	6	1.2	445	1	ACSC_WOOTH	Q07340	moorella th	629	6	1.2	489	1	PEN3_ADEMT	O10439	mouse adeno
557	6	1.2	445	1	ALCP_BACP3	P30145	bacillus ps	630	6	1.2	490	1	DHAB_BACSU	P71016	bacillus su
558	6	1.2	445	1	TBB_LEIME	P21148	leishmania	631	6	1.2	490	1	RBL_CYLSN	P24673	cyllindroche
559	6	1.2	448	1	GNTF_BACLI	P46832	bacillus li	632	6	1.2	490	1	RBL_ODOSI	P02930	escherichia
560	6	1.2	448	1	GNTF_BACSU	P12012	bacillus su	633	6	1.2	491	1	ACHE_BOVIN	P02715	bos taurus
561	6	1.2	449	1	OCF6_HUMAN	O03052	homo sapien	634	6	1.2	491	1	CPH1_CHICK	P05180	gallus gall
562	6	1.2	449	1	OCF6_MOUSE	P21952	mus musculu	635	6	1.2	491	1	CPH2_CHICK	P20678	gallus gall
563	6	1.2	450	1	G6P1_BACSU	P80860	bacillus su	636	6	1.2	491	1	Y084_MYCTU	O53209	mycobacteri
564	6	1.2	450	1	VIME_CARAV	P46673	carassius a	637	6	1.2	492	1	CP53_PIG	Q02390	sus scrofa
565	6	1.2	450	1	YJ59_YEAST	P47048	saccharomyc	638	6	1.2	492	1	FLIC_SALTU	P06179	salmonella
566	6	1.2	451	1	OCF6_RAT	P20267	rattus norv	639	6	1.2	493	1	ACHE_MOUSE	O04844	homo sapien
567	6	1.2	452	1	TRPC_ECOLI	P00909	escherichia	640	6	1.2	493	1	ACHE_MOUSE	P20782	mus musculu
568	6	1.2	452	1	TRPC_SALTU	P00910	salmonella	641	6	1.2	493	1	FLIC_SALPA	P06178	salmonella
569	6	1.2	454	1	APY_SOLTU	P80595	solanum tub	642	6	1.2	494	1	ACHE_RAT	P09660	rattus norv
570	6	1.2	454	1	PR11_SCHPO	O14215	schizosacch	643	6	1.2	495	1	TOLC_ECOLI	P33293	escherichia
571	6	1.2	454	1	PUCG_RHOSH	P95656	rhodovulum	644	6	1.2	496	1	BAF1_KIDNA	P33293	kluyveromyc
572	6	1.2	455	1	VIME_CYPCA	O92155	cyprinus ca	645	6	1.2	498	1	CPB5_ARATH	P65784	arabidopsis
573	6	1.2	457	1	PM4H_CAEEL	P90925	caenorhabdi	646	6	1.2	498	1	POLG_DEN19	P27909	dengue viru
574	6	1.2	458	1	DG17_DICDI	P11467	dicyosteli	647	6	1.2	499	1	MYV1N_AQUAE	O67658	aquifex aeo
575	6	1.2	459	1	PUCG_RHOSH	Q02443	rhodobacter	648	6	1.2	500	1	FLJB_SALAE	P52615	salmonella
576	6	1.2	459	1	RBL_CALSH	P48687	calyptrosph	649	6	1.2	500	1	PTGI_BOVIN	O29626	bos taurus
577	6	1.2	459	1	YAI2_STRCO	O86635	streptomyces	650	6	1.2	500	1	PTGI_HUMAN	O16647	homo sapien
578	6	1.2	460	1	DCED_RHIME	P13632	rhizobium m	651	6	1.2	501	1	DIDH_PEA	P33023	pisum sativ
579	6	1.2	460	1	VL2_HPV44	O80918	human papil	652	6	1.2	502	1	K2M3_SHEEP	O59160	ovis aries
580	6	1.2	460	1	VL2_HPV55	O80939	human papil	653	6	1.2	503	1	OOXA_AGR74	Q59160	agrobacteri
581	6	1.2	461	1	EF1A_ARTSA	P02929	artemia sal	654	6	1.2	504	1	FLIC_SALMU	P06177	salmonella
582	6	1.2	461	1	PR58_XENLA	P46670	xenopus lae	655	6	1.2	504	1	SAH2_DROME	P50245	drosophila
583	6	1.2	462	1	TUBE_DROME	P22812	drosophila	656	6	1.2	504	1	TLR1_DROME	P30974	drosophila
584	6	1.2	465	1	EXSH_RHIME	O33680	rhizobium m	657	6	1.2	505	1	C762_SOLME	P37122	solanum mel
585	6	1.2	465	1	SYN_ECOLI	P17242	escherichia	658	6	1.2	505	1	FLJB_SALTU	P52616	salmonella
586	6	1.2	465	1	YB57_YEAST	P38310	saccharomyc	659	6	1.2	509	1	YMY4_CAEEL	Q10051	caenorhabdi
587	6	1.2	466	1	CF1A_DROME	P16241	drosophila	660	6	1.2	510	1	FKH_DROME	P14734	drosophila
588	6	1.2	467	1	C15Y_CANTR	P79024	candida tro	661	6	1.2	512	1	DDDH_SCHPO	O00087	schizosacch
589	6	1.2	468	1	GLNA_AZOB	P10583	azospirillum	662	6	1.2	512	1	K2C5_XENLA	P15501	autographa
590	6	1.2	468	1	GLNA_ECOLI	P06611	escherichia	663	6	1.2	512	1	VP67_APVAC	P15501	autographa
591	6	1.2	468	1	GLNA_SALTU	P06601	salmonella	664	6	1.2	513	1	TYR8_SCHPO	O05427	schizosacch
592	6	1.2	468	1	GLNA_THIFE	P07804	thiobacilli	665	6	1.2	513	1	TYR8_SCHPO	O05427	schizosacch
593	6	1.2	469	1	GLNA_PROVU	P28786	proteus vul	666	6	1.2	513	1	Y073_TREPA	O83112	treponema p
594	6	1.2	469	1	PPA5_KLUFA	P52289	kluyveromyc	667	6	1.2	514	1	VS14_TREBA	P28329	trypanosoma
595	6	1.2	470	1	YMR7_YEAST	O04371	saccharomyc	668	6	1.2	517	1	YJTI_HAEIN	P44744	haemophilus
596	6	1.2	472	1	GLNA_HAEIN	P43794	haemophilus	669	6	1.2	517	1	CP53_ASPNG	P17549	aspergillus
597	6	1.2	472	1	IFR2_HUMAN	P09913	homo sapien	670	6	1.2	517	1	EOG1_YEAST	P32474	saccharomyc
598	6	1.2	473	1	LCB1_CRIGR	O54695	cricetulus	671	6	1.2	518	1	YX23_CAEEL	Q11194	caenorhabdi
599	6	1.2	473	1	LCB1_HUMAN	O15659	homo sapien	672	6	1.2	519	1	ATPB_NEUCR	P23704	neurospora
600	6	1.2	473	1	LCB1_MOUSE	O35704	mus musculu	673	6	1.2	519	1	CP5K_CANNA	O12589	candida mal
601	6	1.2	474	1	YPC2_CAEEL	P11179	caenorhabdi	674	6	1.2	520	1	TM11_ECOLI	P10484	escherichia
602	6	1.2	475	1	PERI_MOUSE	P15531	mus musculu	675	6	1.2	520	1	TR14_FUSSP	Q16162	fusarium sp
603	6	1.2	476	1	YGI2_YEAST	P53001	saccharomyc	676	6	1.2	520	1	YVAV_SCHPO	O14079	schizosacch
604	6	1.2	479	1	DEAD_BACSU	P42305	bacillus su	677	6	1.2	522	1	YVAV_SCHPO	Q27517	caenorhabdi
605	6	1.2	479	1	HYFD_ECOLI	P77416	escherichia	678	6	1.2	522	1	ABP1_SCHPO	P49777	schizosacch
606	6	1.2	480	1	PYR5_HUMAN	P11172	homo sapien	679	6	1.2	524	1	CK13_YEAST	P39662	saccharomyc
607	6	1.2	480	1	TPSB_ASPNG	O00017	aspergillus	680	6	1.2	524	1	CP5F_CANTR	P39608	candida tro
608	6	1.2	482	1	ODB2_BOVIN	P11181	bos taurus	681	6	1.2	524	1	GLPK_HUMAN	P33189	homo sapien
609	6	1.2	482	1	ODB2_HUMAN	P11182	homo sapien	682	6	1.2	524	1	GLPK_MOUSE	O64516	mus musculu
610	6	1.2	482	1	ODB2_MOUSE	P53395	mus musculu	683	6	1.2	524	1	DEP2_RAT	O64516	mus musculu
611	6	1.2	485	1	YC11_KLEPN	O48457	klebsiella	684	6	1.2	526	1	DEP2_CAEEL	O19196	caenorhabdi
612	6	1.2	486	1	RBL1_RHOSH	P27997	rhodospirillum	685	6	1.2	526	1	MS51_YEAST	P33559	saccharomyc
613	6	1.2	486	1	RBLC_ALCEU	P09657	alcaligenes	686	6	1.2	526	1	COXA_RHIME	P27330	rhizobium m
614	6	1.2	486	1	RBLP_ALCEU	P42721	alcaligenes	687	6	1.2	527	1	TAGH_BACSU	P42954	bacillus su
615	6	1.2	486	1	RBL_RHIME	P56689	rhizobium m	688	6	1.2	527	1	TF65_XENLA	O04865	xenopus lae
616	6	1.2	487	1	CP80_BERST	P47195	berberis st	689	6	1.2	528	1	NIFK_FRALV	O57118	frankia aln
617	6	1.2	488	1	FLIC_SALCH	P06176	salmonella	690	6	1.2	529	1	UDB9_MACFA	O02663	macaca fasc


```

983 6 1.2 1132 1 NTU1_YEAST P53114 saccharomyc
984 6 1.2 1135 1 RB12_PAT 055081 rattus norv
985 6 1.2 1136 1 ADDP_BACSU P23477 bacillus su
986 6 1.2 1167 1 AGE1_CAEEL Q94125 caenorhabd1
987 6 1.2 1184 1 POL2_GPELV P18474 grapevine f
988 6 1.2 1194 1 DPOL_VZVD P09252 varicella-2
989 6 1.2 1195 1 METH_SYNY3 055786 synchocyst
990 6 1.2 1199 1 N121_PAT P52591 rattus norv
991 6 1.2 1209 1 DNB1_HSYEB P28932 equine herp
992 6 1.2 1209 1 THR_DROME P42286 drosophila
993 6 1.2 1220 1 DPOL_HSYEB P28858 equine herp
994 6 1.2 1226 1 POLG_DENIM P17763 dengue viru
995 6 1.2 1248 1 TOPG_SULAC 008582 sulfolobus
996 6 1.2 1250 1 VFAL_ECOLI P45508 escherichia
997 6 1.2 1251 1 RBP2_PLAAB 000799 plasmodium
998 6 1.2 1267 1 VL3_REOVD P17378 reovirus (t
999 6 1.2 1267 1 VL3_REOVL P17376 reovirus (t
1000 6 1.2 1273 1 MYS3_YEAST P36006 saccharomyc

```

ALIGNMENTS

```

RESULT 1
LCRV_YERPE STANDARD; PRT; 326 AA.
ID LCRV_YERPE
AC P21206;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN
V).
GN LCRV.
OS Yersinia pestis.
OG Plasmid pCD1.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM:
RC MEDLINE: 90008806.
RA Price S.B., Leung K.Y., Barve S.S., Straley S.C.;
RT "Molecular analysis of lcrGVH, the V antigen operon of Yersinia
RT pests.";
RT J. Bacteriol. 171:5646-5653(1989).
CC -1- FUNCTION: POSSIBLY INVOLVED IN CA(2+) REGULATION OF YOP
CC EXPRESSION, WHICH INCLUDES THE EXPORT PROCESS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M26405; AAA27641.1; ALT_SEQ.
CC DR PIR: B33601; B33601.
CC KW Plasmid; Antigen; Virulence.
CC SEQUENCE 326 AA; 37226 MW; 54FB8209E032F3F4 CRC64;

```

```

Query Match 36.7%; Score 191; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 1e-175;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 331 LKIVYSNMHNGARSKLRELAELTRELKITYIAQEIKKHLSSSGTINIHDKSINLMD 330
DB 136 LKIVYSNMHNGARSKLRELAELTRELKITYIAQEIKKHLSSSGTINIHDKSINLMD 195
QY 391 KNLGYTDEEIFRASAAYKLEKMPOTTIOVDSEKKIYSIKDFLGSEKRTGALGNLKN 450

```

```

DB 196 KNLGYTDEEIFRASAAYKLEKMPOTTIOVDSEKKIYSIKDFLGSEKRTGALGNLKN 255
QY 451 SYSYNKDNNELSHFATCSDKSRPLNDVYSQKTTQJSDITSRPNISAIEALNRFIOYDSV 510
DB 256 SYSYNKDNNELSHFATCSDKSRPLNDVYSQKTTQJSDITSRPNISAIEALNRFIOYDSV 315
QY 511 MORLDPTSGK 521
DB 316 MORLDPTSGK 326

```

RESULT 2

```

CAP1_YERPE STANDARD; PRT; 170 AA.
ID CAP1_YERPE
AC P26948;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE P1 CAPSULE ANTIGEN PRECURSOR.
GN CAP1.
OS Yersinia pestis.
OG Plasmid pFra.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE: 91099503.
RA Galinov E.E., Smirnov O.Y., Karlishhev A.V., Volkovoy K.I.,
RA Denezuyuk A.I., Nazimov I.V., Rubtsov K.S., Abramov V.M.,
RA Dalvadyanz S.M., Zay'yalov V.P.;
RT "Nucleotide sequence of the Yersinia pestis gene encoding F1 antigen
RT and the primary structure of the protein. Putative T and B cell
RT epitopes.";
RT FEBS Lett. 277:230-232(1990).
CC -1- SUBCELLULAR LOCATION: CAPSULE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X61996; CAA43966.1; -
CC DR PIR: S13008; S13008.
CC KW Plasmid; Signal.
FT SIGNAL 1 21
FT CHAIN 22 170 POTENTIAL.
FT DOMAIN 100 150 CONTAINS POTENTIAL ANTIGENIC DETERMINANTS
FT THAT MAY STIMULATE T-CELLS.
SQ SEQUENCE 170 AA; 17666 MW; 9AC87796A0BA67D9 CRC64;

```

```

Query Match 32.6%; Score 170; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 8.6e-156;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 24 MKRISSVIAIALFGTATNAADLTASTATATVLPARTTLTKRGAGATTINDNINIDT 83
DB 1 MKRISSVIAIALFGTATNAADLTASTATATVLPARTTLTKRGAGATTINDNINIDT 60
QY 84 ELAVGTLLIGGYKTGTSVNFETDAAGDPMYLTFTSODGNHQTFTKYIGKDSRPFDIS 143
DB 61 ELAVGTLLIGGYKTGTSVNFETDAAGDPMYLTFTSODGNHQTFTKYIGKDSRPFDIS 120
QY 144 PKVNGENLVGDVYVLAATGSGDFFVRSIGSKGKLAAGKTTDAVTVTSNQ 193
DB 121 PKVNGENLVGDVYVLAATGSGDFFVRSIGSKGKLAAGKTTDAVTVTSNQ 170

```

```

RESULT 3
LCRV_YERPE

```

```

ID  LCRV_YERPS  STANDARD:  PRT:  326 AA.
AC  P23994;
DT  01-MAR-1992 (Rel. 21, Created)
DT  01-MAR-1992 (Rel. 21, Last sequence update)
DT  15-FEB-2000 (Rel. 39, Last annotation update)
DE  VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN
DE  V).
GN  LCRV.
OS  Yersinia pseudotuberculosis.
OG  Plasmid pIB1.
OC  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC  Yersinia.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-YPIII;
RX  MEDLINE: 91154114.
RA  Bergman T., Haakansson S., Forsberg A., Norlander L., Macellaro A.,
RA  Beekman A., Boelín I., Wolf-Watz H.
RT  "Analysis of the V antigen lcrGVH-yopBD operon of Yersinia
RT  pseudotuberculosis: evidence for a regulatory role of lcrV and
RT  lcrV."
RL  J. Bacteriol. 173:1607-1616(1991).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE: 97162308.
RA  Roggenkamp A., Geiger A.M., Leitritz L., Kessler A., Heesemann J.;
RT  "Passive immunity to infection with Yersinia spp. mediated by anti-
RT  recombinant V antigen is dependent on polymorphism of V antigen."
RL  Infect. Immun. 65:446-451(1997).
CC  -1- FUNCTION: INVOLVED IN CA(2+) REGULATION OF YOP EXPRESSION, WHICH
CC  INCLUDES THE EXPORT PROCESS.
CC  -1- SUBCELLULAR LOCATION: SECRETED.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; M57893; AAA27645.1; -
DR  EMBL; X96802; CAA65594.1; -
DR  PIR; B37314; B37314.
KW  Plasmid; Antigen; Virulence.
SQ  SEQUENCE 326 AA; 37336 MW; 2FD945DAFF48C06 CRC64;

Query Match 15.7%; Score 82; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 5.3e-71;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  196 MIRAEONPHATEDELEKRVQQLTGHGSSVLEBYQVLYKKNIDISITVDPKRSSEVFA 255
DB  1 MIRAEONPHATEDELEKRVQQLTGHGSSVLEBYQVLYKKNIDISITVDPKRSSEVFA 60
OY  256 NRVITDIEFLKRIIAYFLPED 277
DB  61 NRVITDIEFLKRIIAYFLPED 82

RESULT 4
MOX2_HUMAN  STANDARD:  PRT:  303 AA.
AC  P50222;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  15-JUL-1998 (Rel. 36, Last annotation update)
DE  HOMEBOX PROTEIN MOX-2 (GROWTH ARREST-SPECIFIC HOMEBOX).
GN  MOX2 OR MOX2 OR GAX.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE-EMBRYO;
RX  MEDLINE: 95331791.
RA  Grigoriou M., Kastirnakl M.-C., Modl W., Theodorakis K., Mankoo B.,
RA  Pachnis V., Karagogeos D.;
RT  "Isolation of the human MOX2 homeobox gene and localization to
RT  chromosome 7p22.1-p21.3."
RL  Genomics 26:550-555(1995).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE-HEART;
RX  MEDLINE: 95229154.
RA  Lepage D.F., Walsh K.;
RT  "Molecular cloning and localization of the human MOX2 gene to 7p21."
RL  Genomics 24:535-540(1994).
CC  -1- FUNCTION: ROLE IN MESODERM INDUCTION AND ITS EARLIEST REGIONAL
CC  SPECIFICATION, SOMITOGENESIS, AND MYOGENIC AND SCLEROTOMAL
CC  DIFFERENTIATION. MAY HAVE A REGULATORY ROLE WHEN QUIESCENT
CC  VASCULAR SMOOTH MUSCLE CELLS REENTER THE CELL CYCLE (BY
CC  SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC  -1- TISSUE SPECIFICITY: EMBRYO AND PLACENTA.
CC  -1- SIMILARITY: WITH OTHER HOMEBOX PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; X82629; CAA57949.1; -
DR  EMBL; L36328; AAB58497.1; -
DR  HSSP; P02833; ISAN.
DR  MIM; 600535; -
DR  PFM; PF00046; homeobox.1.
DR  PRINTS; PR00024; HOMEBOX.
DR  PROSITE; PS00027; HOMEBOX.1; 1.
DR  PROSITE; PS50071; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Nuclear protein; Developmental protein.
FT  DOMAIN 42 47
FT  DOMAIN 68 79
FT  DOMAIN 80 85
FT  DNA_BIND 186 245
FT  CONFLICT 58 58 HOMEBOX.
FT  CONFLICT 79 79 MISSING (IN REF. 2).
SQ  SEQUENCE 303 AA; 33457 MW; 809ADE0CD090023D CRC64;

Query Match 2.1%; Score 11; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  2 GHHHHHHHHH 12
DB  67 GHHHHHHHHH 77

RESULT 5
MOX2_MOUSE  STANDARD:  PRT:  303 AA.
AC  P32443;
DT  01-OCT-1993 (Rel. 27, Created)
DT  01-OCT-1993 (Rel. 27, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last annotation update)
DE  HOMEBOX PROTEIN MOX-2 OR GAX.
GN  MOX2 OR MOX2 OR GAX.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN  [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE: 93201999.
RA Candia A.F., Hu J., Crosby J., Lalley P.A., Noden D., Nadeau J.H.,
RT Wright C.V.E.;
RT "Mox-1 and Mox-2 define a novel homeobox gene subfamily and are
RT differentially expressed during early mesodermal patterning in mouse
RT embryos."
RT Development 116:1123-1136(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94232829.
RA Candia A.F., Kovalik J.-P., Wright C.V.E.;
RT "Amino acid sequence of Mox-2 and comparison to its Xenopus and rat
RT homologs."
RT Nucleic Acids Res. 21:4982-4982(1993).
RN [3]
RP SEQUENCE OF 1-11 FROM N.A.
RX MEDLINE: 95349593.
RA Andres V., Fisher S., Wearsch P., Walsh K.;
RT "Regulation of Gax homeobox gene transcription by a combination of
RT positive factors including myocyte-specific enhancer factor 2."
RT Mol. Cell. Biol. 15:4272-4281(1995).
CC -1- FUNCTION: ROLE IN MESODERM INDUCTION AND ITS EARLIEST REGIONAL
CC SPECIFICATION, SOMITOGENESIS, AND MYOGENIC AND SCLEROTOMAL
CC DIFFERENTIATION. MAY HAVE A REGULATORY ROLE WHEN QUIESCENT
CC VASCULAR SMOOTH MUSCLE CELLS REENTER THE CELL CYCLE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- DEVELOPMENTAL STAGE: IT IS NOT EXPRESSED BEFORE 8-8.5 DAYS POST
CC COITUM. AT 8-8.5 D.P.C. IT IS FOUND ON THE ENTIRE EPITHELIUM OF
CC THE SOMITE. AT 9.5 D.P.C. ITS EXPRESSION IS RESTRICTED TO THE
CC SCLEROTOME. AT 10.5 D.P.C. IT IS FOUND IN SCLEROTOMALLY DERIVED
CC CELLS INCLUDING THE VERTEBRAL AND COSTAL PRECURSORS.
CC -1- SIMILARITY: WITH OTHER HOMEBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: 216406; CAA7899.1; -
DR EMBL: S79168; CAB33758.1; -
DR PIR: S41779; S41779.
DR HSSP: P02833; ISAN.
DR MGD: MGI:103219; MEOX2.
DR PFAM: PF00046; homeobox. 1.
DR PRINTS: PR00024; HOMEBOX.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS0071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.
FT DOMAIN 42 47 POLY-SER.
FT DOMAIN 68 79 POLY-HIS.
FT DOMAIN 80 85 POLY-GLN.
FT DOMAIN 63 85 GLN/HIS-RICH (OPA-REPEAT).
FT DNA_BIND 186 245 HOMEBOX.
FT SEQUENCE 303 AA; 33506 MW; 41BD05FC39AA4427 CRC64;

```

Query Match 2.1%; Score 11; DB 1; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0079;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 DB 67 GHHHHHHHHH 77

RESULT 6
 ID MOX2_RAT
 AC P39020;
 STANDARD; PRT; 303 AA.

```

DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HOMEBOX PROTEIN MOX-2 (GROWTH ARREST-SPECIFIC HOMEBOX).
GN MEOX2 OR MOX2 OR MOX-2 OR GAX.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-AORTA;
RX MEDLINE: 93268321.
RA Gorski D.H., Lepage D.F., Patel C.V., Copeland N.G., Jenkins N.A.,
RT Walsh K.;
RT "Molecular cloning of a diverged homeobox gene that is rapidly down-
RT regulated during the G0/G1 transition in vascular smooth muscle
RT cells."
RT Mol. Cell. Biol. 13:3722-3733(1993).
RN [2]
RP REVISIONS.
RA Walsh K.;
RT Submitted (MAR-1993) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ROLE IN MESODERM INDUCTION AND ITS EARLIEST REGIONAL
CC SPECIFICATION, SOMITOGENESIS, AND MYOGENIC AND SCLEROTOMAL
CC DIFFERENTIATION. MAY HAVE A REGULATORY ROLE WHEN QUIESCENT
CC VASCULAR SMOOTH MUSCLE CELLS REENTER THE CELL CYCLE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- TISSUE SPECIFICITY: AORTA AND HEART. ALSO DETECTED IN LUNG AND
CC KIDNEY.
CC -1- INDUCTION: RAPIDLY AND TRANSIENTLY DOWN-REGULATED DURING THE
CC TRANSITION FROM G0 TO G1 INDUCED BY MITOGEN STIMULATION.
CC -1- SIMILARITY: WITH OTHER HOMEBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: Z17223; CAA78931.1; -
DR PIR: A48130; A48130.
DR HSSP: P02833; ISAN.
DR PFAM: PF00046; homeobox. 1.
DR PRINTS: PR00024; HOMEBOX.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS0071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.
FT DOMAIN 42 47 POLY-SER.
FT DOMAIN 68 79 POLY-HIS.
FT DOMAIN 80 85 POLY-GLN.
FT DOMAIN 64 85 GLN/HIS-RICH (OPA-REPEAT).
FT DNA_BIND 186 245 HOMEBOX.
FT SEQUENCE 303 AA; 33605 MW; 7776642AEFA3A2E8 CRC64;

```

Query Match 2.1%; Score 11; DB 1; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0079;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 DB 67 GHHHHHHHHH 77

RESULT 7
 ID OTX1_HUMAN
 AC P32242;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE HOMEOBOX PROTEIN OTX1.
 GN OTX1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 93327763.
 RA Simeone A., Acampora D., Mallamaci A., Stornaiuolo A., D'Apice M.R.,
 Nigro V., Boncinelli E.;
 RT "A vertebrate gene related to orthodenticle contains a homeodomain of
 the bicoid class and demarcates anterior neuroectoderm in the
 gastrulating mouse embryo."
 RL EMBL J. 12:2735-2747(1993).
 CC -1- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
 THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS):
 5'-TCTATGCC-3'.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN.
 CC -1- DEVELOPMENTAL STAGE: EMBRYO.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 DR HSSP: P06601; 1FJL.
 DR TRANSFAC: T02079; -.
 DR MIM: 600036; -.
 DR PFAM: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEOBOX.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS50071; HOMEOBOX_2; 1.
 KW Homeobox: DNA-binding; Developmental protein; Nuclear protein.
 FT DNA_BIND 38 97 HOMEOBOX.
 FT DOMAIN 275 301 HIS-RICH.
 SO SEQUENCE 354 AA; 37327 MW; E32C1E47464BDAA CRC64;

Query Match 2.1%; Score 11; DB 1; Length 354;
 Best Local Similarity 100.0%; Pred. No. 0.0091;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 DB 291 GHHHHHHHHH 301

RESULT 8
 OTX1_MOUSE STANDARD: PRT; 355 AA.
 AC P80205;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HOMEOBOX PROTEIN OTX1.
 GN OTX1 OR OTX-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RP SEQUENCE FROM N.A.
 RX STRAIN-C57BL/6;
 RX MEDLINE: 93327763.
 RA Simeone A., Acampora D., Mallamaci A., Stornaiuolo A., D'Apice M.R.,
 Nigro V., Boncinelli E.;
 RT "A vertebrate gene related to orthodenticle contains a homeodomain of
 the bicoid class and demarcates anterior neuroectoderm in the
 gastrulating mouse embryo."
 RL EMBL J. 12:2735-2747(1993).
 CC -1- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
 THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS):
 5'-TCTATGCC-3'.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN.
 CC -1- DEVELOPMENTAL STAGE: EMBRYO.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 DR HSSP: P06601; 1FJL.
 DR TRANSFAC: T02079; -.
 DR MIM: 600036; -.
 DR PFAM: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEOBOX.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS50071; HOMEOBOX_2; 1.
 KW Homeobox: DNA-binding; Developmental protein; Nuclear protein.
 FT DNA_BIND 38 97 HOMEOBOX.
 FT DOMAIN 275 301 HIS-RICH.
 SO SEQUENCE 354 AA; 37327 MW; E32C1E47464BDAA CRC64;

CC -1- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
 THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS):
 5'-TCTATGCC-3'.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: BRAIN: RESTRICTED REGIONS OF THE DEVELOPING
 ROSTRAL BRAIN INCLUDING THE PRESUMPTIVE CEREBRAL CORTEX AND
 OLFACTORY BULBS; EXPRESSED IN THE DEVELOPING OLFACTORY, AUDICULAR
 AND OCULAR SYSTEMS, INCLUDING THE COVERING OF THE OPTIC NERVE.
 CC -1- DEVELOPMENTAL STAGE: EMBRYO.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X68883; CAA48754.1; -.
 DR PIR: S35345; S35345.
 DR HSSP: P06601; 1FJL.
 DR TRANSFAC: T02080; -.
 DR MGD: MGI:97450; OTX1.
 DR PFAM: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEOBOX.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS50071; HOMEOBOX_2; 1.
 KW Homeobox: DNA-binding; Developmental protein; Nuclear protein.
 FT DNA_BIND 38 97 HOMEOBOX.
 FT DOMAIN 275 302 HIS-RICH.
 SO SEQUENCE 355 AA; 37531 MW; DA5136D9DAFC948F CRC64;

Query Match 2.1%; Score 11; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.0091;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 DB 291 GHHHHHHHHH 301

RESULT 9
 OTX1_RAT STANDARD: PRT; 355 AA.
 AC Q63410; Q64203;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HOMEOBOX PROTEIN OTX1.
 GN OTX1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RP SEQUENCE FROM N.A.
 RX TISSUE-BRAIN;
 RX MEDLINE: 95016961.
 RA Frantz G.D., Weimann J.M., Levin M.E., McConnell S.K.;
 RT "Otx1 and Otx2 define layers and regions in developing cerebral
 cortex and cerebellum."
 RL J. Neurosci. 14:5725-5740(1994).
 CC -1- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
 THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS):
 5'-TCTATGCC-3'.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: BRAIN: RESTRICTED REGIONS OF THE DEVELOPING
 ROSTRAL BRAIN INCLUDING THE PRESUMPTIVE CEREBRAL CORTEX AND
 OLFACTORY BULBS; EXPRESSED IN THE DEVELOPING OLFACTORY, AUDICULAR
 AND OCULAR SYSTEMS, INCLUDING THE COVERING OF THE OPTIC NERVE.
 CC -1- DEVELOPMENTAL STAGE: EMBRYO.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 DR EMBL: X68883; CAA48754.1; -.
 DR PIR: S35345; S35345.
 DR HSSP: P06601; 1FJL.
 DR TRANSFAC: T02080; -.
 DR MGD: MGI:97450; OTX1.
 DR PFAM: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEOBOX.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS50071; HOMEOBOX_2; 1.
 KW Homeobox: DNA-binding; Developmental protein; Nuclear protein.
 FT DNA_BIND 38 97 HOMEOBOX.
 FT DOMAIN 275 302 HIS-RICH.
 SO SEQUENCE 355 AA; 37531 MW; DA5136D9DAFC948F CRC64;

CC THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS): 5'-
 CC TCTATCCC-3' (BY SIMILARITY). MAY PLAY A ROLE IN THE SPECIFICATION
 CC OR DIFFERENTIATION OF NEURONS IN THE DEEP LAYERS OF THE CEREBRAL
 CC CORTX, AND ALSO IN CEREBELLAR REGIONALIZATION DURING EARLY
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: BRAIN. RESTRICTED REGIONS OF THE DEVELOPING
 CC ROSTRAL BRAIN INCLUDING THE PRESUMPTIVE CEREBRAL CORTX AND
 CC OLFACTORY BULBS; EXPRESSED IN THE DEVELOPING OLFACTORY, AUDICULAR
 CC AND OCULAR SYSTEMS, INCLUDING THE COVERING OF THE OPTIC NERVE (BY
 CC SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT THE FOREBRAIN AND
 CC MIDBRAIN DURING DEVELOPMENT. AND IN ADDITION IS ALSO SEEN IN
 CC DISCRETE SPATIAL AND TEMPORAL DOMAINS IN THE DEVELOPING CEREBRAL
 CC CORTX AND CEREBELLUM. CONFINED TO A SUBPOPULATION OF NEURONS IN
 CC LAYERS 5 AND 6 WITHIN THE ADULT CEREBRAL CORTX AND DURING
 CC DEVELOPMENT EXPRESSION IS HIGH IN THE PROGENITORS OF THESE DEEP-
 CC LAYER CELLS. EXPRESSED IN THE DEVELOPING CEREBELLUM IN SPATIALLY
 CC RESTRICTED REGIONS OF THE EXTERNAL GRANULAR LAYER.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L33602; AA53557.1; -;
 CC EMBL: S81924; CAB3993.1; -;
 CC DR HSP: P06601; IPIJL.
 CC DR PFAM: PF00046; homeobox; 1.
 CC DR PRINTS: PR00024; HOMEOBOX.
 CC DR PROSITE: PS00027; HOMEOBOX_1; 1.
 CC DR PROSITE: PS50071; HOMEOBOX_2; 1.
 CC KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 CC FT DNA_BIND 38 97 HOMEOBOX.
 CC FT DOMAIN 275 302 HIS-RICH.
 CC FT CONFLICT 219 219 T -> A (IN REF. 2).
 CC FT CONFLICT 229 229 R -> G (IN REF. 2).
 CC SO SEQUENCE 355 AA; 37602 MW; C875871723D0B876 CRC64;

Query Match 2.1%; Score 11; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.0091;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHHNNHHNNH 12
 DB 291 GHHNNHHNNH 301

RESULT 10
 ID TY11_HUMAN STANDARD; PRT; 414 AA.
 AC P25490; Q14935;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE TRANSCRIPTIONAL REPRESSOR PROTEIN Y11 (Y11 AND YANG 1) (YY-1) (DELTA
 DE TRANSCRIPTION FACTOR) (NF-E1).
 DE Y11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE; 92005716.
 RA Shi Y., Seto E., Chang L.-S., Shenk T.;
 RT "Transcriptional repression by Y11, a human G1-Kruppel-related
 RT protein, and relief of repression by adenovirus E1a protein.";

RL Cell 67:377-386(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-FORESKIN;
 RX MEDLINE; 92052179.
 RA Park K., Atchison M.;
 RT "Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta),
 RT that binds to the immunoglobulin kappa 3' enhancer and the
 RT immunoglobulin heavy-chain mu E1 site.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9804-9808(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Whitson R.H., Huang T., Dang J., Itakura K.;
 RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE; 98152931.
 RA McNeil S., Guo B., Stein J.L., Lian J.B., Bushmeyer S., Seto E.,
 RA Atchison M.L., Pennan S., van Wijnen A.J., Stein G.S.;
 RT "Targeting of the Y11 transcription factor to the nucleolus and the
 RT nuclear matrix in situ: the C-terminus is a principal determinant for
 RT nuclear trafficking.";
 RL J. Cell. Biochem. 68:500-510(1998).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 304-414.
 RX MEDLINE; 97098436.
 RA Houbavly H.B., Ushveva A., Shenk T., Burley S.K.;
 RT "Cocrystal structure of Y11 bound to the adeno-associated virus P5
 RT Initiator.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13577-13582(1996).
 RN [6]
 RP STRUCTURE BY NMR OF 353-379.
 RX MEDLINE; 98308000.
 RA Viles J.H., Patel S.U., Mitchell J.B.O., Moody C.M., Justice D.E.,
 RA Updegrink J., Doyle P.M., Harris C.J., Sadler P.J., Thornton J.M.;
 RT "Design, synthesis and structure of a zinc finger with an artificial
 RT beta-turn.";
 RL J. Mol. Biol. 279:973-986(1998).
 CC -1- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
 CC POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
 CC VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION
 CC START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
 CC DIFFERENTIATION. THE FUNCTION OF Y11 AS AN ACTIVATOR OR A
 CC REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. FOR
 CC EXAMPLE IT ACTS AS AN ACTIVATOR IN ITS PRESENCE.
 CC PROTEIN BUT AS AN ACTIVATOR IN ITS PRESENCE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
 CC -1- SIMILARITY: BELONGS TO THE KRUPPEL SUBFAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M77698; AA59467.1; -;
 CC EMBL: M76541; AA58926.1; -;
 CC EMBL: Z14077; CAH78435.1; -;
 CC PIR: A40350; A40350.
 CC DR PDB: 1UBD; 23-DEC-96.
 CC DR PDB: 1ZNM; 01-APR-98.
 CC DR TRANSPAC: T00915; -;
 CC DR MTM: 600013; -;
 CC DR PFAM: PF00096; Zf-C2H2; 4.
 CC DR PRINTS: PR00048; ZINC_FINGER.
 CC DR PROSITE: PS00028; ZINC_FINGER_C2H2; 4.
 CC KW Transcription regulation; Repressor; Activator; Nuclear protein;
 CC Zinc-finger; Metal-binding; DNA-binding; Repeat; 3D-structure.
 FT DOMAIN 43 53 ASF/GUT-RICH (ACIDIC).
 FT DOMAIN 54 69 GLY-RICH.

```

FT DOMAIN 70 80 POLY-HIS
FT DOMAIN 159 170 GLY/SER-RICH.
FT FT 296 407 4 C2H2-TYPE ZINC-FINGERS.
FT FT 296 330 C2H2-TYPE.
FT FT 325 347 C2H2-TYPE.
FT FT 353 377 C2H2-TYPE.
FT FT 383 407 C2H2-TYPE.
FT FT 257 341 INVOLVED IN NUCLEAR MATRIX ASSOCIATION.
FT FT 333 371 INVOLVED IN REPRESSION OF ACTIVATED
TRANSCRIPTION.
FT FT 371 397 INVOLVED IN MASKING TRANSCRIPTION
DOMAIN.
FT FT 65 65 H -> R (IN REF. 2).
FT FT 196 196 G -> R (IN REF. 1).
FT FT 414 AA: 44712 MW; 058C05A0AD2D04E6 CRC64;
SQ SEQUENCE

Query Match 2.1%; Score 11; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
Db 69 GHHHHHHHHH 79

RESULT 11
TY1_MOUSE STANDARD; PRT; 414 AA.
ID TY1_MOUSE
AC 000899;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TRANSCRIPTIONAL REPRESSOR PROTEIN YY1 (YY1 AND YANG 1) (YY-1) (DELTA
DE TRANSCRIPTION FACTOR) (NF-E1) (UCR-MOTIF DNA-BINDING PROTEIN).
GN YY1 OR UCRBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 92107191.
RA Flanagan J.R., Becker K.G., Enlist D.L., Gleason S.L., Driggers P.H.,
RA Levi B.-Z., Appella E., Ozato R.;
RT "Cloning of a negative transcription factor that binds to the
RT upstream conserved region of Moloney murine leukemia virus.";
RT Mol. Cell. Biol. 12:38-44(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE: 93296177.
RA Safran G., Perry R.P.;
RT "Characterization of the mouse gene that encodes the delta/YY1/NF-
RT E1/OCBP transcription factor.";
RT Proc. Natl. Acad. Sci. U.S.A. 90:5559-5563(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92052178.
RA Haribaran N., Kelley D.E., Perry R.P.;
RT "Delta, a transcription factor that binds to downstream elements in
RT several polymerase II promoters, is a functionally versatile zinc
RT finger protein.";
RT Proc. Natl. Acad. Sci. U.S.A. 88:9799-9803(1991).
CC -1- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
CC POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
CC VIRAL GENES. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
CC DIFFERENTIATION. THE FUNCTION OF YY1 AS AN ACTIVATOR OR A
CC REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. BINDS TO
CC THE UPSTREAM CONSERVED REGION (UCR) (5'-GCCCATTTT-3') OF MOLONEY
CC MURINE LEUKEMIA VIRUS (MDLV).
CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-

```

```

CC CC FINGER PROTEINS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sib.ch).
CC -----
DR EMBL: M73963; AAA40522.1; -
DR EMBL: L13968; AAA40477.1; -
DR EMBL: L13969; AAA40477.1; JOINED.
DR EMBL: L13965; AAA40477.1; JOINED.
DR EMBL: L13966; AAA40477.1; JOINED.
DR EMBL: L13967; AAA40477.1; JOINED.
DR EMBL: M74580; AAA37521.1; -
DR HSSP: P23480; 10BD.
DR MGD: MGI:99150; YY1.
DR PFAM: PF00096; zf-C2H2; 4.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 4.
DR PROSITE: PS00048; ZINC_FINGER.
KM Transcription regulation: Repressor; Activator; Nuclear protein.
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT FT DOMAIN 43 53 GLY/SER-RICH.
FT FT DOMAIN 54 70 POLY-HIS.
FT FT DOMAIN 71 82 GLY-RICH.
FT FT DOMAIN 161 170 POLY-HIS.
FT FT DOMAIN 286 407 GLY/SER-RICH.
FT FT ZN_FING 296 320 4 C2H2-TYPE ZINC-FINGERS.
FT FT ZN_FING 325 347 C2H2-TYPE.
FT FT ZN_FING 353 377 C2H2-TYPE.
FT FT ZN_FING 383 407 C2H2-TYPE.
FT FT DOMAIN 257 341 INVOLVED IN NUCLEAR MATRIX ASSOCIATION
FT FT DOMAIN 333 371 (BY SIMILARITY).
FT FT DOMAIN 371 397 INVOLVED IN REPRESSION OF ACTIVATED
TRANSCRIPTION (BY SIMILARITY).
FT FT DOMAIN 371 397 INVOLVED IN MASKING TRANSCRIPTION
DOMAIN (BY SIMILARITY).
FT FT CONFLICT 219 219 F -> S (IN REF. 3).
FT FT CONFLICT 375 375 R -> G (IN REF. 3).
FT FT SEQUENCE 414 AA: 44717 MW; C01237828B984F9 CRC64;

Query Match 2.1%; Score 11; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
Db 70 GHHHHHHHHH 80

RESULT 12
ZIC3_MOUSE STANDARD; PRT; 466 AA.
ID ZIC3_MOUSE
AC 062521;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ZINC FINGER PROTEIN ZIC3 (ZINC FINGER PROTEIN OF THE CEREBELLUM 3).
GN ZIC3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 96132843.
RA Aruga J., Nagai T., Tokuyama T., Hayashizaki Y., Okazaki Y.,
RA Chapman V.N., Mikoshiba K.;
RT "The mouse zic gene family. Homologues of the Drosophila pair-rule
RT gene odd-paired.";

```



```

CC BL Chem. 271.1043-1047(1996).
CC -1- FUNCTION: PROBABLY FUNCTIONS AS A TRANSCRIPTION FACTOR IN THE
CC EARLIEST STAGES OF THE LEFT-RIGHT (LR) BODY AXIS FORMATION (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: CNS. A HIGH LEVEL EXPRESSION IS SEEN IN THE
CC CEREBELLUM.
CC -1- SIMILARITY: TO THE GLI-RELATED GROUP OF C2H2-TYPE ZINC-FINGERS
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D70849; BAA1116.1; -.
DR HSP: P08047; ISP2.
DR MGD: MGI:106676; ZIC3.
DR PRAM: PFO0096; zf-C2H2; 5.
DR PRINTS: PFO0048; ZINC-FINGER.
DR PROSITE: PS00028; ZINC-FINGER_C2H2; 3.
DR Transcription regulation; zinc-finger; Metal-binding; DNA-binding;
KM Repeat; Nuclear protein.
FT DOMAIN 46 55 POLY-ALA.
FT DOMAIN 87 96 POLY-HIS.
FT ZN_FING 294 321 C2H2-TYPE (ATYPICAL).
FT ZN_FING 327 351 C2H2-TYPE.
FT ZN_FING 357 381 C2H2-TYPE.
FT ZN_FING 387 409 C2H2-TYPE.
SQ SEQUENCE 466 AA; 50644 MW; 089749DF726F069B CRC64;

Query Match 2.1%; Score 11; DB 1; Length 466;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHHHHHHHHH 12
Db 86 GHHHHHHHHH 96

RESULT 13
ZIC3_HUMAN STANDARD; PRT; 467 AA.
ID ZIC3_HUMAN
AC 060481;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ZINC FINGER PROTEIN ZIC3 (ZINC FINGER PROTEIN OF THE CEREBELLUM 3).
GN ZIC3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A. AND VARIANT HTX1 MET-233.
RX MEDLINE; 98016414.
RA Gebbia M., Ferrero G.B., Pilla G., Baszi M.T., Aylsworth A.S.,
RA Pennen-Splitt M., Bird L.M., Bamforth J.S., Burn J., Schlesinger D.,
RA Nelson D.L., Casey B.;
RT "X-linked situs abnormalities result from mutations in ZIC3.";
RL Nat. Genet. 17:305-308(1997).
CC -1- FUNCTION: PROBABLY FUNCTIONS AS A TRANSCRIPTION FACTOR IN THE
CC EARLIEST STAGES OF THE LEFT-RIGHT (LR) BODY AXIS FORMATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DISEASE: DEFECTS IN ZIC3 ARE THE CAUSE OF X-LINKED VISCERAL
CC HETEROTAXY (HTX1), A DISEASE CHARACTERIZED BY CONGENITAL HEART
CC DISEASE AND ALTERATIONS OF VISCERAL SITES, INCLUDING ASPLENIA OR
CC POLYSPLLENIA, SYMMETRIC LIVER, INTESTINAL MALROTATION, AND ABNORMAL
CC LUNG LOBATION.
CC -1- SIMILARITY: TO THE GLI-RELATED GROUP OF C2H2-TYPE ZINC-FINGERS

```

```

CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF028706; AAC05594.1; -.
DR MIM: 306955; -.
DR PRAM: PFO0096; zf-C2H2; 5.
DR PRINTS: PFO0048; ZINC-FINGER.
DR PROSITE: PS00028; ZINC-FINGER_C2H2; 3.
DR Transcription regulation; zinc-finger; Metal-binding; DNA-binding;
KM Repeat; Nuclear protein; Disease mutation.
FT DOMAIN 46 55 POLY-ALA.
FT DOMAIN 87 97 POLY-HIS.
FT ZN_FING 295 322 C2H2-TYPE (ATYPICAL).
FT ZN_FING 328 352 C2H2-TYPE.
FT ZN_FING 358 382 C2H2-TYPE.
FT ZN_FING 388 410 C2H2-TYPE.
FT VARIANT 323 323 T->M (IN HTX1).
SQ SEQUENCE 467 AA; 50569 MW; 3150CF13C0679568 CRC64;

Query Match 2.1%; Score 11; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHHHHHHHHH 12
Db 86 GHHHHHHHHH 96

RESULT 14
SNF1_CANGA STANDARD; PRT; 611 AA.
ID SNF1_CANGA
AC 000372;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CARBON CATABOLITE DEREPRESSING PROTEIN KINASE (EC 2.7.1.-).
GN SNF1.
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC anamorphic Saccharomycetales; Candida.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-NCCLS84;
RX MEDLINE; 97101049.
RA Petter R., Kwon-Chung K.J.;
RT "Disruption of the SNF1 gene abolishes trehalose utilization in the
RT pathogenic yeast Candida glabrata.";
RL Infect. Immun. 64:5269-5273(1996).
CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
CC PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```



```

DR EMBL: L78130; AAB48642.1; -.
DR HSSP: P24941; 1A01.
DR PFAM: PF00069; PKinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
FT DOMAIN 6 17 POLY-HIS.
FT NP_BIND 39 290 PROTEIN_KINASE.
FT BINDING 45 53 ATP (BY SIMILARITY).
FT BINDING 68 68 ATP (BY SIMILARITY).
FT ACT_SITE 161 161 BY SIMILARITY.
FT MOD_RES 194 194 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 611 AA; 70049 MW; 89E17812A490CCD0 CRC64;

```

```

Query Match          2.1%; Score 11; DB 1; Length 611;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 HHHHHHHHHS 13
   |||||
DB 8 HHHHHHHHHS 18

```

```

RESULT 15
SNF1_CANTR
ID SNF1_CANTR STANDARD; PRT; 619 AA.
AC 094168;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CARBON CATABOLITE DEREPRESSING PROTEIN KINASE (EC 2.7.1.-).
GN SNF1
OS Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC anamorphic Saccharomycetales; Candida.
RN [1]
RP SEQUENCE FROM N.A.
RA Kanai T., Ogawa K., Ueda M., Tanaka A.;
RT "Genetic evaluation of the function of SNF1 in Candida tropicalis.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
CC PROTEIN SNF4. COULD PHOSPHORYLATES CAP8 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB024535; BAA75889.1; -.
DR PFAM: PF00069; PKinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
FT DOMAIN 17 29 POLY-HIS.
FT NP_BIND 52 303 PROTEIN_KINASE.
FT BINDING 58 66 ATP (BY SIMILARITY).
FT BINDING 81 81 ATP (BY SIMILARITY).
FT ACT_SITE 174 174 BY SIMILARITY.
FT MOD_RES 207 207 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 619 AA; 70323 MW; 0FCF1FC3DCE706D7 CRC64;

```

```

Query Match          2.1%; Score 11; DB 1; Length 619;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 HHHHHHHHHS 13
   |||||
DB 20 HHHHHHHHHS 30

```

```

Search completed: August 22, 2000, 17:47:11
Job time: 310 sec

```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:41:11 ; Search time 46.55 seconds
(without alignments)
776.006 Million cell updates/sec

Title: US-08-699-716a-2
Perfect score: 521
Sequence: 1 MGNHHHHHHHSSGHIDD.....RFIOKYDSVMORLLDITSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 225878 seqs, 69334122 residues

Word size : 0

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

SPREMBL_12:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	243	46.6	326	2	068697	068697 yersinia pe	
2	104	20.0	334	2	056853	056853 yersinia en	
3	104	20.0	334	2	056896	056896 yersinia en	
4	96	18.4	324	2	056846	056846 yersinia en	
5	96	18.4	324	2	056879	056879 yersinia en	
6	93	17.9	325	2	056880	056880 yersinia en	
7	68	13.1	324	2	087495	087495 yersinia en	
8	67	12.9	324	2	056882	056882 yersinia en	
9	3.3	168	4	016256	homo sapien	016256 homo sapien	
10	17	3.3	294	2	030527	pseudomonas	030527 pseudomonas
11	12	2.3	516	13	042186	brachydanio	042186 brachydanio
12	11	2.1	230	4	075263	homo sapien	075263 homo sapien
13	11	2.1	420	3	042980	schizosacch	042980 schizosacch
14	11	2.1	427	3	P78807	schizosacch	P78807 schizosacch
15	11	2.1	467	13	09YGC6	xenopus lae	09YGC6 xenopus lae
16	11	2.1	911	5	097423	drosophila	097423 drosophila
17	11	2.1	1180	5	024163	plasmodia	024163 plasmodia
18	10	1.9	53	14	055244	gallus galli	055244 gallus galli
19	10	1.9	202	13	Q90691	gallus galli	Q90691 gallus galli

20	10	1.9	204	6	028555 ovis aries
21	10	1.9	309	5	018751 caenorhabd
22	10	1.9	321	10	023891 oryza sativ
23	10	1.9	391	10	023038 arabidopsis
24	10	1.9	407	10	0920A3 arabidopsis
25	10	1.9	443	4	014979 homo sapien
26	10	1.9	465	11	P70512 rattus norv
27	10	1.9	465	11	008755 mus musculu
28	10	1.9	491	11	008755 mus musculu
29	10	1.9	551	11	070369 mus musculu
30	10	1.9	551	11	055187 mus musculu
31	10	1.9	558	4	000257 mus musculu
32	10	1.9	588	12	066043 simian herp
33	10	1.9	604	13	042472 xenopus lae
34	10	1.9	716	5	097216 leishmania
35	10	1.9	791	11	070368 mus musculu
36	10	1.9	884	5	061543 drosophila
37	10	1.9	905	5	026344 aplysia sp.
38	10	1.9	977	11	P97838 rattus norv
39	10	1.9	980	4	095387 homo sapien
40	10	1.9	1100	4	094875 homo sapien
41	10	1.9	1172	10	064474 arabidopsis
42	10	1.9	1196	11	035413 rattus norv
43	10	1.9	1212	5	024523 drosophila
44	10	1.9	1273	4	060316 homo sapien
45	10	1.9	1291	5	077261 drosophila
46	10	1.9	1622	5	065550 cryptospori
47	9	1.7	44	12	065825 undentifile
48	9	1.7	136	11	008657 rattus norv
49	9	1.7	216	13	09W7C7 oryzias lat
50	9	1.7	229	10	065124 dicentra ex
51	9	1.7	311	13	090370 coturnix co
52	9	1.7	311	13	090888 gallus galli
53	9	1.7	324	13	029682 brachydanio
54	9	1.7	348	10	094457 arabidopsis
55	9	1.7	356	13	073679 brachydanio
56	9	1.7	394	5	09YXQ3 anopheles g
57	9	1.7	497	13	09YIB7 xenopus lae
58	9	1.7	501	13	093487 xenopus lae
59	9	1.7	501	13	091689 xenopus lae
60	9	1.7	509	5	094888 drosophila
61	9	1.7	509	5	09YOH8 drosophila
62	9	1.7	515	11	054949 mus musculu
63	9	1.7	522	5	09XTM5 drosophila
64	9	1.7	522	5	09XTK9 drosophila
65	9	1.7	533	4	095409 homo sapien
66	9	1.7	627	4	094825 homo sapien
67	9	1.7	678	5	094736 stomoxys ca
68	9	1.7	693	5	096680 schizosacch
69	9	1.7	794	3	042890 schizosacch
70	9	1.7	808	4	015083 homo sapien
71	9	1.7	1116	3	099247 saccharomyc
72	9	1.7	1555	4	094770 homo sapien
73	9	1.7	1560	4	060275 homo sapien
74	9	1.7	2243	4	09Y5T3 homo sapien
75	9	1.7	2250	4	09Y5T2 homo sapien
76	9	1.7	2288	11	09WUB8 mus musculu
77	9	1.7	2295	11	09WUT2 mus musculu
78	9	1.7	2353	4	095802 homo sapien
79	9	1.7	83	5	020689 caenorhabd
80	8	1.5	102	5	094189 caenorhabd
81	8	1.5	115	1	09YAJ2 aeropyrum p
82	8	1.5	140	5	026056 plasmodium
83	8	1.5	183	5	020741 caenorhabd
84	8	1.5	208	2	084365 chlamydia t
85	8	1.5	216	11	061099 mus musculu
86	8	1.5	216	11	064279 mus musculu
87	8	1.5	250	11	P97832 rattus norv
88	8	1.5	250	5	P90780 caenorhabd
89	8	1.5	274	10	09XGS2 glycine max
90	8	1.5	286	13	057342 coturnix co
91	8	1.5	286	13	042290 gallus galli
92	8	1.5	292	13	091294 rana catesb

93	1.5	323	4	09Y503	09Y593	homo sapien	166	7	1.3	290	5	045469	045469	caenorhabd
94	1.5	326	13	093336	093336	xenopus lae	167	7	1.3	294	2	P94395	P94395	bacillus su
95	1.5	358	12	010418	010418	human cytom	168	7	1.3	295	10	022807	022807	arabidopsi
96	1.5	358	12	069215	069215	human cytom	169	7	1.3	295	10	092V85	092V85	arabidopsi
97	1.5	370	5	024150	024150	dtrosophila	170	7	1.3	305	5	P91314	P91314	caenorhabd
98	1.5	401	3	006178	006178	saccharomye	171	7	1.3	306	11	035449	035449	mus muscu
99	1.5	409	3	012128	012128	saccharomye	172	7	1.3	314	2	025229	025229	helicobacte
100	1.5	423	12	010417	010417	human cytom	173	7	1.3	314	5	021760	021760	arabidopsi
101	1.5	484	11	092206	092206	mus muscu	174	7	1.3	314	10	092UY7	092UY7	arabidopsi
102	1.5	485	4	095948	095948	homo sapien	175	7	1.3	322	10	064452	064452	nicotiana p
103	1.5	559	5	09XUL9	09XUL9	caenorhabd	176	7	1.3	329	2	09XC70	09XC70	streptomye
104	1.5	576	5	024180	024180	dtrosophila	177	7	1.3	342	10	065004	065004	lycopersico
105	1.5	684	12	P87888	P87888	human cytom	178	7	1.3	342	10	042658	042658	caenorhabd
106	1.5	704	5	09XTN4	09XTN4	dtrosophila	179	7	1.3	343	5	001575	001575	caenorhabd
107	1.5	895	4	09Y2X9	09Y2X9	homo sapien	180	7	1.3	348	8	099921	099921	cyprinella
108	1.5	1133	5	095050	095050	tetrahymena	181	7	1.3	348	8	099922	099922	cyprinella
109	1.5	1172	4	09Y4F2	09Y4F2	homo sapien	182	7	1.3	373	11	008349	008349	arabidopsi
110	1.5	1203	5	022484	022484	caenorhabd	183	7	1.3	377	11	008349	008349	arabidopsi
111	1.5	1441	5	096957	096957	dtrosophila	184	7	1.3	384	10	082136	082136	thermotoga
112	1.5	1557	5	096652	096652	dtrosophila	185	7	1.3	394	2	P96877	P96877	mycobacteri
113	1.5	1597	5	061346	061346	dtrosophila	186	7	1.3	406	2	047280	047280	escherichia
114	1.5	1929	5	093637	093637	caenorhabd	187	7	1.3	411	2	P78701	P78701	metarhizium
115	1.5	2416	5	P91667	P91667	dtrosophila	188	7	1.3	415	2	092361	092361	streptomye
116	1.5	5327	5	076891	076891	dtrosophila	189	7	1.3	425	5	076831	076831	caenorhabd
117	1.3	47	12	09WH80	09WH80	chimpanzee	190	7	1.3	434	13	023513	023513	arabidopsi
118	1.3	77	5	020630	020630	caenorhabd	191	7	1.3	444	2	050193	050193	xenopus lae
119	1.3	83	2	P76094	P76094	notophthalm	192	7	1.3	446	6	097727	097727	frankia sp.
120	1.3	127	5	092385	092385	synecococc	193	7	1.3	451	10	082690	082690	leptocinclis
121	1.3	136	5	093827	093827	caenorhabd	194	7	1.3	454	13	073627	073627	saccharopol
122	1.3	142	2	09X781	09X781	streptomye	195	7	1.3	456	2	054101	054101	anolis caro
123	1.3	146	2	092F22	092F22	mycobacteri	196	7	1.3	458	2	048873	048873	lactobacill
124	1.3	148	2	034913	034913	enterococcu	197	7	1.3	466	2	032907	032907	mycobacteri
125	1.3	149	2	P97026	P97026	homo sapien	198	7	1.3	470	2	09ZDM0	09ZDM0	ricketsia
126	1.3	154	4	09Y3C5	09Y3C5	homo sapien	199	7	1.3	470	10	004317	004317	arabidopsi
127	1.3	158	5	094577	094577	helioicidari	200	7	1.3	483	2	054179	054179	streptomye
128	1.3	158	5	046248	046248	dtrosophila	201	7	1.3	499	5	044658	044658	caenorhabd
129	1.3	159	5	046258	046258	dtrosophila	202	7	1.3	504	12	081961	081961	human papil
130	1.3	159	12	09YMT6	09YMT6	lymantria d	203	7	1.3	513	2	032561	032561	escherichia
131	1.3	182	1	059259	059259	pyrococcus	204	7	1.3	513	10	040029	040029	hordium vul
132	1.3	185	10	000837	000837	leishmania	205	7	1.3	513	10	040030	040030	hordium vul
133	1.3	187	10	049684	049684	arabidopsi	206	7	1.3	513	10	081992	081992	hordium vul
134	1.3	192	5	046232	046232	dtrosophila	207	7	1.3	529	10	080567	080567	arabidopsi
135	1.3	192	5	046260	046260	dtrosophila	208	7	1.3	532	2	069975	069975	streptomye
136	1.3	193	5	046262	046262	dtrosophila	209	7	1.3	533	6	097623	097623	canis famli
137	1.3	195	5	046262	046262	dtrosophila	210	7	1.3	533	13	09YGX2	09YGX2	gallus gall
138	1.3	196	5	046234	046234	dtrosophila	211	7	1.3	539	12	09YQSO	09YQSO	human calic
139	1.3	196	5	046234	046234	dtrosophila	212	7	1.3	539	12	09WRZ1	09WRZ1	normal-lik
140	1.3	198	5	046238	046238	dtrosophila	213	7	1.3	539	12	09WRZ0	09WRZ0	normal-lik
141	1.3	198	5	046240	046240	dtrosophila	214	7	1.3	539	12	09WRZ0	09WRZ0	normal-lik
142	1.3	198	5	046244	046244	dtrosophila	215	7	1.3	539	12	09WRZ0	09WRZ0	normal-lik
143	1.3	198	5	046246	046246	dtrosophila	216	7	1.3	539	12	09WRZ0	09WRZ0	normal-lik
144	1.3	199	5	050412	050412	magnetospir	217	7	1.3	539	12	09WRZ0	09WRZ0	normal-lik
145	1.3	201	12	084640	084640	parametium	218	7	1.3	539	12	09WRZ0	09WRZ0	normal-lik
146	1.3	209	12	084640	084640	parametium	219	7	1.3	539	12	09WRZ0	09WRZ0	normal-lik
147	1.3	211	4	075333	075333	homo sapien	220	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
148	1.3	215	4	096004	096004	homo sapien	221	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
149	1.3	219	12	055579	055579	leucania se	222	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
150	1.3	235	5	076647	076647	caenorhabd	223	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
151	1.3	236	5	054779	054779	streptococc	224	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
152	1.3	244	10	064461	064461	brassica na	225	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
153	1.3	253	3	093949	093949	candida alb	226	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
154	1.3	253	3	094052	094052	caenorhabd	227	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
155	1.3	256	4	013742	013742	homo sapien	228	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
156	1.3	256	4	015747	015747	dictyosteli	229	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
157	1.3	259	10	066901	066901	agutifex aeo	230	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
158	1.3	259	10	09ZTC8	09ZTC8	arabidopsi	231	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
159	1.3	263	5	020288	020288	caenorhabd	232	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
160	1.3	263	12	065175	065175	african swi	233	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
161	1.3	273	5	076688	076688	caenorhabd	234	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
162	1.3	276	13	006962	006962	salmonella	235	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
163	1.3	281	2	P79788	P79788	gallus gall	236	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
164	1.3	281	2	067136	067136	agutifex aeo	237	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik
165	1.3	284	2	054187	054187	streptomye	238	7	1.3	540	12	09WRZ0	09WRZ0	normal-lik

239	7	1.3	703	11	061230	Q61230 mus musculu	312	6	1.2	79	2	Q9X214
240	7	1.3	732	4	014586	Q14586 homo sapien	313	6	1.2	79	10	Q42170
241	7	1.3	733	4	028972	Q28972 archaeoglob	314	6	1.2	81	2	054558
242	7	1.3	793	4	09Y2L9	Q9Y2L9 homo sapien	315	6	1.2	84	12	090930
243	7	1.3	794	2	09X212	Q9X212 thermotoga	316	6	1.2	85	6	P79306
244	7	1.3	795	5	09Y1J5	Q9Y1J5 dictyostell	317	6	1.2	85	12	Q9MBE1
245	7	1.3	798	11	088411	Q88411 mus musculu	318	6	1.2	86	2	Q48627
246	7	1.3	799	4	094874	Q94874 homo sapien	319	6	1.2	86	12	Q9YWS4
247	7	1.3	807	12	09YTQ5	Q9YTQ5 ateline her	320	6	1.2	86	12	Q9YWR3
248	7	1.3	808	12	069320	Q69320 gallid heap	321	6	1.2	87	5	001463
249	7	1.3	820	4	000236	Q00236 homo sapien	322	6	1.2	87	12	055317
250	7	1.3	829	5	024183	Q24183 drosophila	323	6	1.2	88	2	066152
251	7	1.3	843	4	09Y2T1	Q9Y2T1 homo sapien	324	6	1.2	88	4	Q94795
252	7	1.3	844	12	041789	Q41789 human immun	325	6	1.2	88	12	Q9YWS5
253	7	1.3	886	13	042588	Q42588 xenopus lae	326	6	1.2	88	12	Q9WCM4
254	7	1.3	892	2	055676	Q55676 synchocyst	327	6	1.2	89	2	P74811
255	7	1.3	904	11	008971	Q08971 mus musculu	328	6	1.2	89	2	Q9ZH04
256	7	1.3	932	3	P89499	P89499 saccharomyc	329	6	1.2	89	2	Q9X008
257	7	1.3	1001	5	022534	Q22534 caenorhabd1	330	6	1.2	89	12	Q78979
258	7	1.3	1002	2	087109	Q87109 serralia ma	331	6	1.2	89	12	069683
259	7	1.3	1041	10	023536	Q23536 arabidopsis	332	6	1.2	90	10	022295
260	7	1.3	1133	5	024034	Q24034 drosophila	333	6	1.2	90	12	069703
261	7	1.3	1133	5	077426	Q77426 drosophila	334	6	1.2	90	12	069652
262	7	1.3	1150	5	094548	Q94548 drosophila	335	6	1.2	90	12	069674
263	7	1.3	1176	12	Q98587	Q98587 paramecium	336	6	1.2	90	12	Q40341
264	7	1.3	1189	4	043604	Q43604 homo sapien	337	6	1.2	90	12	Q9YWS2
265	7	1.3	1199	4	060668	Q60668 homo sapien	338	6	1.2	91	12	069515
266	7	1.3	1199	4	043487	Q43487 homo sapien	339	6	1.2	91	12	069697
267	7	1.3	1289	3	008748	Q08748 saccharomyc	340	6	1.2	91	12	Q76757
268	7	1.3	1314	4	Q12769	Q12769 homo sapien	341	6	1.2	91	12	Q76758
269	7	1.3	1318	12	P90493	P90493 herpes simp	342	6	1.2	91	12	Q76759
270	7	1.3	1361	2	086617	Q86617 streptomyc	343	6	1.2	91	12	Q76760
271	7	1.3	1378	12	Q9WR57	Q9WR57 macaca mula	344	6	1.2	91	12	Q76761
272	7	1.3	1402	11	0960W3	Q960W3 mus musculu	345	6	1.2	91	12	Q76762
273	7	1.3	1465	5	017909	Q17909 caenorhabd1	346	6	1.2	91	12	Q9WCM5
274	7	1.3	1530	4	043241	Q43241 homo sapien	347	6	1.2	91	12	Q9WCM5
275	7	1.3	1531	2	084818	Q84818 Chlamydia t	348	6	1.2	92	2	Q9XCH6
276	7	1.3	1705	2	005179	Q05179 bordetella	349	6	1.2	92	2	Q9XCH6
277	7	1.3	1894	11	064487	Q64487 mus musculu	350	6	1.2	92	5	019331
278	7	1.3	1979	5	065133	Q65133 plasmodium	351	6	1.2	92	12	Q79228
279	7	1.3	2117	2	Q92429	Q92429 streptomyc	352	6	1.2	92	12	Q79228
280	7	1.3	3056	12	065852	Q65852 bean yellow	353	6	1.2	92	12	Q79228
281	7	1.3	3056	12	065892	Q65892 bean yellow	354	6	1.2	92	12	Q79228
282	7	1.3	3072	12	093645	Q93645 clover yell	355	6	1.2	92	12	Q79228
283	7	1.3	3104	2	004846	Q04846 corynebacte	356	6	1.2	92	12	Q79228
284	7	1.3	3345	5	046074	Q46074 drosophila	357	6	1.2	92	12	Q79228
285	7	1.3	3413	2	054593	Q54593 amycolatops	358	6	1.2	92	12	Q79228
286	7	1.3	3729	2	033956	Q33956 streptomyc	359	6	1.2	92	12	Q79228
287	7	1.3	3729	2	033956	Q33956 streptomyc	360	6	1.2	92	12	Q79228
288	6	1.2	16	2	048439	Q48439 klebsiella	361	6	1.2	92	12	Q79228
289	6	1.2	34	5	094583	Q94583 heliothis v	362	6	1.2	92	12	Q79228
290	6	1.2	41	2	006527	Q06527 lactobacill	363	6	1.2	92	12	Q79228
291	6	1.2	54	12	069824	Q69824 human immun	364	6	1.2	92	12	Q79228
292	6	1.2	54	12	069826	Q69826 human immun	365	6	1.2	92	12	Q79228
293	6	1.2	57	12	069114	Q69114 human herpe	366	6	1.2	92	12	Q79228
294	6	1.2	57	12	069122	Q69122 human herpe	367	6	1.2	92	12	Q79228
295	6	1.2	61	12	090268	Q90268 hepatitis b	368	6	1.2	92	12	Q79228
296	6	1.2	62	2	034390	Q34390 bacillus su	369	6	1.2	92	12	Q79228
297	6	1.2	64	10	Q40049	Q40049 hordeum vul	370	6	1.2	92	12	Q79228
298	6	1.2	65	1	Q51794	Q51794 pyrococcus	371	6	1.2	92	12	Q79228
299	6	1.2	66	10	Q22226	Q22226 arabidopsis	372	6	1.2	92	12	Q79228
300	6	1.2	70	11	P71998	P71998 mycobacteri	373	6	1.2	92	12	Q79228
301	6	1.2	70	11	088380	Q88380 mus musculu	374	6	1.2	92	12	Q79228
302	6	1.2	72	2	025973	Q25973 helicobacte	375	6	1.2	92	12	Q79228
303	6	1.2	73	9	092X34	Q92X34 mycobacteri	376	6	1.2	92	12	Q79228
304	6	1.2	76	5	Q23550	Q23550 naegleria f	377	6	1.2	92	12	Q79228
305	6	1.2	76	10	065488	Q65488 arabidopsis	378	6	1.2	92	12	Q79228
306	6	1.2	77	12	Q92X18	Q92X18 helicobacte	379	6	1.2	92	12	Q79228
307	6	1.2	77	12	Q9WC11	Q9WC11 human immun	380	6	1.2	92	12	Q79228
308	6	1.2	78	2	032805	Q32805 cyprinus ca	381	6	1.2	92	12	Q79228
309	6	1.2	78	2	Q9XV55	Q9XV55 caenorhabd1	382	6	1.2	92	12	Q79228
310	6	1.2	78	12	Q70489	Q70489 human immun	383	6	1.2	92	12	Q79228
311	6	1.2	79	2	085256	Q85256 streptococ	384	6	1.2	92	12	Q79228

385	6	1.2	93	12	Q9YWS3	Q9YWS3 human immun	458	6	1.2	112	1	027980	027980 archaeoglob
386	6	1.2	93	12	Q9WCP2	Q9WCP2 human immun	459	6	1.2	112	12	P88144	P88144 human immun
387	6	1.2	93	12	Q9WCL7	Q9WCL7 human immun	460	6	1.2	113	13	042461	042461 squallus aca
388	6	1.2	93	12	Q9WCK8	Q9WCK8 human immun	461	6	1.2	113	13	042462	042462 squallus aca
389	6	1.2	93	12	Q9WCH6	Q9WCH6 human immun	462	6	1.2	113	13	012967	012967 fuigu rubrip
390	6	1.2	94	12	Q97198	Q97198 human immun	463	6	1.2	114	5	046154	046154 plutella xy
391	6	1.2	94	12	Q97204	Q97204 human immun	464	6	1.2	114	5	046155	046155 plutella xy
392	6	1.2	94	12	Q9WCT4	Q9WCT4 human immun	465	6	1.2	115	5	096793	096793 myzus persi
393	6	1.2	94	12	Q9WCS6	Q9WCS6 human immun	466	6	1.2	115	5	017259	017259 brachionus
394	6	1.2	94	12	Q9WCS3	Q9WCS3 human immun	467	6	1.2	115	12	Q9YL39	Q9YL39 human immun
395	6	1.2	94	12	Q9WCR4	Q9WCR4 human immun	468	6	1.2	115	12	Q9YKGS	Q9YKGS human immun
396	6	1.2	94	12	Q9WCH7	Q9WCH7 human immun	469	6	1.2	116	2	P76837	P76837 escherichia
397	6	1.2	94	12	Q9WCH0	Q9WCH0 human immun	470	6	1.2	116	12	067741	067741 human astro
398	6	1.2	94	12	Q9WCG9	Q9WCG9 human immun	471	6	1.2	116	12	067742	067742 human astro
399	6	1.2	95	5	Q24866	Q24866 entamoeba i	472	6	1.2	117	8	020157	020157 beet virus
400	6	1.2	95	9	Q38398	Q38398 bacterioph	473	6	1.2	117	8	09XMS5	09XMS5 tetrahymena
401	6	1.2	95	11	P70644	P70644 rattus norv	474	6	1.2	118	5	019424	019424 caenorhabdi
402	6	1.2	95	12	Q37099	Q37099 human immun	475	6	1.2	118	12	Q92517	Q92517 beet virus
403	6	1.2	95	12	Q9WL21	Q9WL21 human immun	476	6	1.2	120	1	058193	058193 pyrococcus
404	6	1.2	95	12	Q9WL16	Q9WL16 human immun	477	6	1.2	121	2	Q49098	Q49098 arabidopsis
405	6	1.2	95	12	Q9WCS2	Q9WCS2 human immun	478	6	1.2	121	10	023089	023089 arabidopsis
406	6	1.2	95	12	Q9WCS4	Q9WCS4 human immun	479	6	1.2	123	12	Q9YWB1	Q9YWB1 human immun
407	6	1.2	95	12	Q9WCL1	Q9WCL1 human immun	480	6	1.2	124	2	068464	068464 rickettsia
408	6	1.2	95	12	Q9WCH9	Q9WCH9 human immun	481	6	1.2	124	2	068465	068465 rickettsia
409	6	1.2	95	12	Q9W9H1	Q9W9H1 human immun	482	6	1.2	125	2	Q9XD69	Q9XD69 legionella
410	6	1.2	95	13	Q73812	Q73812 morone saxa	483	6	1.2	125	12	P90116	P90116 human immun
411	6	1.2	96	12	Q9WCU1	Q9WCU1 human immun	484	6	1.2	127	2	Q9WYH4	Q9WYH4 thermotoga
412	6	1.2	96	12	Q9WCU0	Q9WCU0 human immun	485	6	1.2	128	2	P94214	P94214 lactobacill
413	6	1.2	96	12	Q9WCS7	Q9WCS7 human immun	486	6	1.2	128	3	059957	059957 agaveus bi
414	6	1.2	96	12	Q9WCP6	Q9WCP6 human immun	487	6	1.2	128	10	Q9XFL7	Q9XFL7 phaseolus v
415	6	1.2	96	12	Q9WCM3	Q9WCM3 human immun	488	6	1.2	131	2	P94342	P94342 burholderi
416	6	1.2	96	12	Q9WCM2	Q9WCM2 human immun	489	6	1.2	132	5	Q9XJ97	Q9XJ97 lampetra fl
417	6	1.2	96	12	Q9WCL8	Q9WCL8 human immun	490	6	1.2	133	2	Q9ZK84	Q9ZK84 helicobacte
418	6	1.2	96	12	Q9WNA3	Q9WNA3 human immun	491	6	1.2	133	3	Q9ZEP9	Q9ZEP9 pseudomonas
419	6	1.2	96	12	Q9W9V2	Q9W9V2 human immun	492	6	1.2	134	12	Q66118	Q66118 cucumber mo
420	6	1.2	98	2	Q31026	Q31026 vibrio chol	493	6	1.2	134	12	Q74951	Q74951 human immun
421	6	1.2	98	5	Q76496	Q76496 heliothis v	494	6	1.2	134	12	P88703	P88703 human immun
422	6	1.2	98	12	Q07958	Q07958 human immun	495	6	1.2	134	12	P90138	P90138 human immun
423	6	1.2	99	12	Q9WCH1	Q9WCH1 human immun	496	6	1.2	136	2	Q07663	Q07663 enterococu
424	6	1.2	100	2	Q84034	Q84034 chlamydia t	497	6	1.2	136	5	Q45207	Q45207 anopheles g
425	6	1.2	103	2	Q49507	Q49507 mycoplasma	498	6	1.2	136	5	P90543	P90543 eubacteri
426	6	1.2	103	12	Q87075	Q87075 pseudorabie	499	6	1.2	137	2	Q9ZIC8	Q9ZIC8 listeria mo
427	6	1.2	103	12	Q9WBF4	Q9WBF4 human immun	500	6	1.2	137	5	Q96334	Q96334 ditrofilaria
428	6	1.2	104	12	Q9WBD8	Q9WBD8 human immun	501	6	1.2	137	5	Q96335	Q96335 brugia mala
429	6	1.2	105	2	Q85812	Q85812 streptococ	502	6	1.2	137	5	Q96336	Q96336 onchocerca
430	6	1.2	105	4	Q15944	Q15944 homo sapien	503	6	1.2	137	9	Q38374	Q38374 lactococcus
431	6	1.2	105	11	Q70155	Q70155 mus musculu	504	6	1.2	139	3	Q74900	Q74900 schistosach
432	6	1.2	106	12	Q9WCT6	Q9WCT6 human immun	505	6	1.2	139	5	P90945	P90945 caenorhabdi
433	6	1.2	107	12	Q90587	Q90587 human immun	506	6	1.2	139	12	Q75346	Q75346 human immun
434	6	1.2	107	12	Q90588	Q90588 human immun	507	6	1.2	139	12	Q9WAE5	Q9WAE5 influenza c
435	6	1.2	107	12	Q90590	Q90590 human immun	508	6	1.2	139	12	Q9WAD8	Q9WAD8 influenza c
436	6	1.2	107	12	Q9YRT3	Q9YRT3 human immun	509	6	1.2	139	12	Q9WAD6	Q9WAD6 influenza c
437	6	1.2	107	12	Q9YRT2	Q9YRT2 human immun	510	6	1.2	139	12	Q9WAD4	Q9WAD4 influenza c
438	6	1.2	107	12	Q9YRT0	Q9YRT0 human immun	511	6	1.2	139	12	Q9W8T3	Q9W8T3 influenza c
439	6	1.2	107	12	Q9YRS9	Q9YRS9 human immun	512	6	1.2	139	12	Q9W8T5	Q9W8T5 influenza c
440	6	1.2	107	12	Q9YRS8	Q9YRS8 human immun	513	6	1.2	139	12	Q9W8L5	Q9W8L5 influenza c
441	6	1.2	107	12	Q9YRS7	Q9YRS7 human immun	514	6	1.2	140	5	Q17821	Q17821 caenorhabdi
442	6	1.2	107	12	Q9YRS5	Q9YRS5 human immun	515	6	1.2	142	2	Q50753	Q50753 borrelia bu
443	6	1.2	107	12	Q9YRS4	Q9YRS4 human immun	516	6	1.2	142	5	Q97167	Q97167 culic pipie
444	6	1.2	107	12	Q9YRS3	Q9YRS3 human immun	517	6	1.2	142	6	Q9XSL4	Q9XSL4 sus scrofa
445	6	1.2	107	12	Q9YRS2	Q9YRS2 human immun	518	6	1.2	142	11	Q61522	Q61522 mus musculu
446	6	1.2	107	12	Q9YRP2	Q9YRP2 human immun	519	6	1.2	143	2	Q48129	Q48129 lactococcus
447	6	1.2	107	12	Q9YRX5	Q9YRX5 human immun	520	6	1.2	143	5	Q20291	Q20291 caenorhabdi
448	6	1.2	107	12	Q9WLD5	Q9WLD5 human immun	521	6	1.2	143	5	Q76223	Q76223 trypanosoma
449	6	1.2	108	5	Q99099	Q99099 onchocerca	522	6	1.2	143	9	Q9ZAX7	Q9ZAX7 bacterioph
450	6	1.2	108	12	Q55902	Q55902 human immun	523	6	1.2	144	5	Q01257	Q01257 bacterioph
451	6	1.2	109	1	Q58646	Q58646 pyrococcus	524	6	1.2	144	9	Q48447	Q48447 bacterioph
452	6	1.2	110	6	Q46627	Q46627 bos taurus	525	6	1.2	145	6	Q18926	Q18926 sus scrofa
453	6	1.2	110	12	Q96712	Q96712 cucumber mo	526	6	1.2	146	1	Q58599	Q58599 methanococ
454	6	1.2	111	12	Q50328	Q50328 escherichia	527	6	1.2	146	2	Q50021	Q50021 mycobacteri
455	6	1.2	111	12	Q39437	Q39437 cucumber mo	528	6	1.2	146	10	Q24024	Q24024 lycopersico
456	6	1.2	111	12	Q86784	Q86784 cucumber mo	529	6	1.2	146	10	Q41514	Q41514 tritpsacum a
457	6	1.2	111	12	Q9WA72	Q9WA72 cucumber mo	530	6	1.2	147	2	Q9WYK3	Q9WYK3 thermotoga

531	6	1.2	149	2	005444	005444 mycobacteri	604	1.2	191	5	026511	026511 schistosoma
532	6	1.2	149	2	09wxf4	09wxf4 thermotoga	605	1.2	191	5	003528	003528 schistosoma
533	6	1.2	150	4	09y4m1	09y4m1 homo sapien	606	1.2	191	5	001372	001372 schistosoma
534	6	1.2	150	5	002581	002581 inciliaria f	607	1.2	192	11	051109	051109 mus musculu
535	6	1.2	150	5	017962	017962 caenorhabdi	608	1.2	193	5	090560	090560 oxytricha f
536	6	1.2	150	12	098778	098778 vesicular s	609	1.2	193	5	094976	094976 oxytricha f
537	6	1.2	150	12	098779	098779 vesicular s	610	1.2	194	2	069747	069747 neisseria l
538	6	1.2	150	12	098780	098780 vesicular s	611	1.2	194	2	086479	086479 streptococ
539	6	1.2	150	12	098790	098790 vesicular s	612	1.2	194	2	096853	096853 schistosoma
540	6	1.2	150	12	098791	098791 vesicular s	613	1.2	194	2	096853	096853 schistosoma
541	6	1.2	150	12	098791	098791 vesicular s	614	1.2	195	6	028171	028171 capra hircu
542	6	1.2	150	12	098787	098787 vesicular s	615	1.2	196	2	028171	028171 capra hircu
543	6	1.2	151	12	098789	098789 vesicular s	616	1.2	197	1	059040	059040 pyrococcus
544	6	1.2	151	5	025283	025283 loligo fort	617	1.2	197	10	041837	041837 zea mays (m
545	6	1.2	152	5	016955	016955 acanthamoeb	618	1.2	198	1	057724	057724 pyrococcus
546	6	1.2	152	5	025314	025314 leishmania	619	1.2	198	4	09y6x8	09y6x8 homo sapien
547	6	1.2	154	5	029686	029686 archaeoglob	620	1.2	199	2	087795	087795 prochloroco
548	6	1.2	154	5	017739	017739 caenorhabdi	621	1.2	199	10	065687	065687 arabidopsis
549	6	1.2	155	10	093828	093828 arabidopsis	622	1.2	200	2	073246	073246 synecocyst
550	6	1.2	157	1	026871	026871 methanobact	623	1.2	200	12	092239	092239 human immun
551	6	1.2	157	2	047443	047443 escherichia	624	1.2	201	4	099935	099935 homo sapien
552	6	1.2	158	1	058766	058766 pyrococcus	625	1.2	201	12	09wkm9	09wkm9 human immun
553	6	1.2	158	2	062377	062377 clostridium	626	1.2	202	5	073378	073378 plasmodium
554	6	1.2	158	10	081775	081775 arabidopsis	627	1.2	202	10	041643	041643 volvox cart
555	6	1.2	160	3	012540	012540 agaricus bl	628	1.2	202	12	071827	071827 human immun
556	6	1.2	161	2	09x6m4	09x6m4 streptococ	629	1.2	203	3	09yrc7	09yrc7 human immun
557	6	1.2	162	2	033056	033056 mycobacteri	630	1.2	203	5	061828	061828 caenorhabdi
558	6	1.2	162	5	061829	061829 caenorhabdi	631	1.2	203	12	070726	070726 human immun
559	6	1.2	162	10	09xw56	09xw56 arabidopsis	632	1.2	204	3	008907	008907 saccharomyc
560	6	1.2	163	2	059307	059307 corneobacte	633	1.2	205	1	028634	028634 archaeoglob
561	6	1.2	166	2	030415	030415 lactococcus	634	1.2	205	12	09w1l7	09w1l7 human immun
562	6	1.2	167	2	067892	067892 aquifex aeo	635	1.2	206	5	061761	061761 caenorhabdi
563	6	1.2	168	1	057996	057996 pyrococcus	636	1.2	206	5	045868	045868 caenorhabdi
564	6	1.2	168	2	067861	067861 aquifex aeo	637	1.2	208	2	083947	083947 trepomema p
565	6	1.2	168	2	046662	046662 brucella ab	638	1.2	208	2	09x612	09x612 streptococ
566	6	1.2	168	5	09xay0	09xay0 pseudomonas	639	1.2	208	5	022156	022156 caenorhabdi
567	6	1.2	169	5	09xw11	09xw11 caenorhabdi	640	1.2	208	8	078939	078939 agrocylbe ae
568	6	1.2	169	13	093343	093343 xenopus lae	641	1.2	208	12	090025	090025 human immun
569	6	1.2	170	2	09wv88	09wv88 pseudomonas	642	1.2	209	2	09xcs2	09xcs2 legionella
570	6	1.2	171	5	061171	061171 entodinium	643	1.2	209	12	085953	085953 human immun
571	6	1.2	171	5	046254	046254 scapitomyza	644	1.2	210	5	025448	025448 musca domes
572	6	1.2	172	2	032418	032418 staphylococ	645	1.2	210	5	062813	062813 papio hamad
573	6	1.2	172	2	076572	076572 escherichia	646	1.2	210	8	078286	078286 palmaria pa
574	6	1.2	172	2	0926N8	0926N8 chlamydia p	647	1.2	210	11	070377	070377 rattus norv
575	6	1.2	172	2	09x879	09x879 streptomyce	648	1.2	210	11	035620	035620 mus musculu
576	6	1.2	173	2	047446	047446 escherichia	649	1.2	210	11	009044	009044 mus musculu
577	6	1.2	173	2	047190	047190 escherichia	650	1.2	211	2	026053	026053 mus musculu
578	6	1.2	173	2	09xav5	09xav5 pseudomonas	651	1.2	211	2	09xj99	09xj99 helicobacte
579	6	1.2	173	5	062574	062574 podocoryne	652	1.2	211	2	09x2x4	09x2x4 bacillus an
580	6	1.2	173	5	061172	061172 entodinium	653	1.2	213	2	025918	025918 helicobacte
581	6	1.2	175	1	028291	028291 archaeoglob	654	1.2	213	2	073135	073135 synecocyst
582	6	1.2	175	2	047449	047449 escherichia	655	1.2	214	2	09xw82	09xw82 streptomyce
583	6	1.2	175	9	003942	003942 bacterioph	656	1.2	215	5	025445	025445 musca domes
584	6	1.2	177	2	092xj9	092xj9 helicobacte	657	1.2	215	5	025450	025450 musca domes
585	6	1.2	177	11	054841	054841 mus musculu	658	1.2	215	12	09y6p7	09y6p7 sweet potat
586	6	1.2	181	1	028254	028254 archaeoglob	659	1.2	216	5	09xw55	09xw55 caenorhabdi
587	6	1.2	183	5	018248	018248 caenorhabdi	660	1.2	216	12	090053	090053 human immun
588	6	1.2	184	5	022026	022026 caenorhabdi	661	1.2	218	5	09xwfo	09xwfo caenorhabdi
589	6	1.2	184	10	080364	080364 arabidopsis	662	1.2	218	11	064401	064401 cricetus
590	6	1.2	184	12	088133	088133 arbididopsis	663	1.2	218	11	060466	060466 cricetus
591	6	1.2	185	13	079857	079857 viana virus	664	1.2	219	5	025514	025514 helicobacte
592	6	1.2	186	2	09wyt4	09wyt4 thermotoga	665	1.2	219	5	017346	017346 caenorhabdi
593	6	1.2	186	5	001589	001589 caenorhabdi	666	1.2	220	2	069465	069465 mycobacteri
594	6	1.2	186	10	09xw58	09xw58 arabidopsis	667	1.2	221	2	034310	034310 bacillus su
595	6	1.2	187	12	09yrm8	09yrm8 human immun	668	1.2	221	2	09w2x2	09w2x2 thermotoga
596	6	1.2	187	12	09yrm8	09yrm8 human immun	669	1.2	221	10	096477	096477 lycopersico
597	6	1.2	188	3	099321	099321 saccharomyc	670	1.2	222	2	086632	086632 streptomyce
598	6	1.2	189	5	023631	023631 cryptospori	671	1.2	222	2	09x622	09x622 bacillus sp
599	6	1.2	190	1	029539	029539 archaeoglob	672	1.2	222	12	082125	082125 viral hemor
600	6	1.2	190	4	046236	046236 drosophila	673	1.2	222	12	084475	084475 viral hemor
601	6	1.2	191	4	094931	094931 homo sapien	674	1.2	223	10	041429	041429 solanum tub
602	6	1.2	191	5	044372	044372 schistosoma	675	1.2	223	12	076769	076769 human immun
603	6	1.2	191	5	017091	017091 caenorhabdi	676	1.2	223	12	076770	076770 human immun

677	6	1.2	223	12	076771	076771 human immun	750	6	1.2	256	6	002756	002756 bos taurus
678	6	1.2	223	12	076772	076772 human immun	751	6	1.2	257	2	031204	031204 pasteurella
679	6	1.2	224	2	066646	066646 aquifex aeo	752	6	1.2	257	2	084195	084195 chlamydia t
680	6	1.2	225	2	092571	092571 zymomonas m	753	6	1.2	257	10	09XG32	09XG32 guillardiella
681	6	1.2	225	10	092033	092033 arabidopsis	754	6	1.2	258	1	029264	029264 archaeoglob
682	6	1.2	226	1	026341	026341 methanobact	755	6	1.2	258	1	09YDB9	09YDB9 aeropyrum p
683	6	1.2	226	2	006319	006319 mycobacteri	756	6	1.2	258	2	056836	056836 amoeba prot
684	6	1.2	226	2	083590	083590 treponema p	757	6	1.2	258	2	09X414	09X414 chlamydiophl
685	6	1.2	226	5	024082	024082 diatrophila	758	6	1.2	258	10	064595	064595 arabidopsis
686	6	1.2	226	12	098686	098686 simian cyto	759	6	1.2	259	10	082798	082798 arabidopsis
687	6	1.2	227	2	09WM59	09WM59 pseudomonas	760	6	1.2	259	12	098687	098687 simian cyto
688	6	1.2	227	8	09XLO7	09XLO7 falco pereg	761	6	1.2	260	2	025113	025113 helicobacte
689	6	1.2	230	2	085466	085466 bacillus ce	762	6	1.2	260	2	09XAB5	09XAB5 bacillus ce
690	6	1.2	230	4	017503	017503 caenorhabdi	763	6	1.2	262	2	09WZP6	09WZP6 thermotoga
691	6	1.2	231	4	095372	095372 homo sapien	764	6	1.2	262	5	097313	097313 plasmidum
692	6	1.2	231	10	09ZWS9	09ZWS9 arabidopsis	765	6	1.2	262	12	092547	092547 epizootic h
693	6	1.2	231	11	09WRL7	09WRL7 mus musculu	766	6	1.2	262	14	009195	009195 unidentified
694	6	1.2	232	2	053668	053668 streptomyce	767	6	1.2	263	12	084599	084599 chlamydia t
695	6	1.2	232	2	055942	055942 synechocyst	768	6	1.2	263	12	071100	071100 bovine aden
696	6	1.2	233	2	074284	074284 synechocyst	769	6	1.2	263	12	09WDX5	09WDX5 human immun
697	6	1.2	233	5	096361	096361 hyphartria	770	6	1.2	265	2	007791	007791 mycobacteri
698	6	1.2	234	2	056180	056180 synechococ	771	6	1.2	265	5	016807	016807 hirudo medi
699	6	1.2	235	2	0928L9	0928L9 chlamydia p	772	6	1.2	265	11	054830	054830 mus musculu
700	6	1.2	235	5	002442	002442 globodera r	773	6	1.2	266	2	09XRG2	09XRG2 streptomyce
701	6	1.2	235	7	098224	098224 chlamydomon	774	6	1.2	266	11	035334	035334 mus musculu
702	6	1.2	235	8	033368	033368 chlamydomon	775	6	1.2	267	8	048006	048006 thomomys ma
703	6	1.2	236	1	028882	028882 archaeoglob	776	6	1.2	267	8	035178	035178 neacoms sp
704	6	1.2	237	2	09X830	09X830 streptomyce	777	6	1.2	268	4	075806	075806 homo sapien
705	6	1.2	238	12	090337	090337 saquaro cac	778	6	1.2	268	14	009194	009194 unidentified
706	6	1.2	239	1	029017	029017 archaeoglob	779	6	1.2	269	10	039360	039360 brassica na
707	6	1.2	239	1	09YAS7	09YAS7 aeropyrum p	780	6	1.2	270	2	050653	050653 paracoccus
708	6	1.2	239	5	090778	090778 caenorhabdi	781	6	1.2	270	3	073854	073854 synechocyst
709	6	1.2	240	5	096178	096178 plasmidum	782	6	1.2	270	3	094105	094105 pleurotus d
710	6	1.2	241	2	032827	032827 legionella	783	6	1.2	270	10	024141	024141 nicotiana t
711	6	1.2	241	2	084728	084728 chlamydia t	784	6	1.2	271	2	005879	005879 mycobacteri
712	6	1.2	241	10	022849	022849 arabidopsis	785	6	1.2	271	10	004633	004633 arabidopsis
713	6	1.2	242	5	032767	032767 legionella	786	6	1.2	272	10	022806	022806 arabidopsis
714	6	1.2	242	5	018238	018238 caenorhabdi	787	6	1.2	273	10	092740	092740 arabidopsis
715	6	1.2	242	5	096136	096136 plasmidum	788	6	1.2	274	10	081395	081395 actinifilum
716	6	1.2	242	6	028075	028075 bos taurus	789	6	1.2	274	12	089048	089048 vesicular s
717	6	1.2	243	1	057997	057997 pyrococcus	790	6	1.2	275	10	064700	064700 arabidopsis
718	6	1.2	243	2	09ZFU0	09ZFU0 streptococ	791	6	1.2	277	1	028441	028441 archaeoglob
719	6	1.2	243	9	080208	080208 methanobact	792	6	1.2	277	12	09WEG0	09WEG0 human immun
720	6	1.2	244	5	095601	095601 rhizobium e	793	6	1.2	277	12	09WDX7	09WDX7 human immun
721	6	1.2	244	5	016421	016421 caenorhabdi	794	6	1.2	278	2	053104	053104 rhodobacter
722	6	1.2	245	5	018039	018039 caenorhabdi	795	6	1.2	278	2	092811	092811 chlamydia p
723	6	1.2	246	13	090811	090811 gallus gall	796	6	1.2	278	5	061831	061831 caenorhabdi
724	6	1.2	247	2	053566	053566 haemophilus	797	6	1.2	279	1	09YBPI	09YBPI aeropyrum p
725	6	1.2	247	5	020284	020284 caenorhabdi	798	6	1.2	279	12	09WEG6	09WEG6 human immun
726	6	1.2	247	10	092914	092914 actinidia c	799	6	1.2	280	2	09WYGI	09WYGI thermotoga
727	6	1.2	247	11	070140	070140 mus musculu	800	6	1.2	280	10	041602	041602 triticum du
728	6	1.2	248	3	043033	043033 schizosacch	801	6	1.2	281	2	054359	054359 streptomyce
729	6	1.2	248	5	045450	045450 caenorhabdi	802	6	1.2	281	6	077728	077728 ovis aries
730	6	1.2	248	12	09WTF6	09WTF6 human immun	803	6	1.2	281	8	09XWTF	09XWTF elephas max
731	6	1.2	249	5	046357	046357 caenorhabdi	804	6	1.2	282	12	057164	057164 spinach lat
732	6	1.2	250	2	09Z1A2	09Z1A2 clostridium	805	6	1.2	283	2	085831	085831 sphingomona
733	6	1.2	250	2	09Z7V0	09Z7V0 chlamydia p	806	6	1.2	284	2	0928Y3	0928Y3 chlamydia p
734	6	1.2	250	5	090779	090779 caenorhabdi	807	6	1.2	284	4	014874	014874 homo sapien
735	6	1.2	250	5	017721	017721 caenorhabdi	808	6	1.2	284	4	09XWPF	09XWPF elephas max
736	6	1.2	251	2	032768	032768 legionella	809	6	1.2	285	1	09YAB5	09YAB5 aeropyrum p
737	6	1.2	252	2	051905	051905 proteus mir	810	6	1.2	285	2	074637	074637 synechocyst
738	6	1.2	252	3	094663	094663 schizosacch	811	6	1.2	285	5	094553	094553 dugesia tlg
739	6	1.2	252	10	09ZU30	09ZU30 arabidopsis	812	6	1.2	285	5	045492	045492 caenorhabdi
740	6	1.2	253	2	067894	067894 aquifex aeo	813	6	1.2	285	8	020957	020957 cyrhonectr
741	6	1.2	253	2	083296	083296 treponema p	814	6	1.2	285	8	09XWTF	09XWTF elephas max
742	6	1.2	253	5	026869	026869 typanosoma	815	6	1.2	285	10	082166	082166 arabidopsis
743	6	1.2	253	11	09WY65	09WY65 rattus norv	816	6	1.2	286	2	024958	024958 helicobacte
744	6	1.2	253	2	054169	054169 streptomyce	817	6	1.2	286	12	088977	088977 venezuelan
745	6	1.2	255	2	09WZ47	09WZ47 thermotoga	818	6	1.2	287	10	048781	048781 arabidopsis
746	6	1.2	255	2	026335	026335 methanobact	819	6	1.2	288	3	094366	094366 schizosacch
747	6	1.2	256	2	031618	031618 bacillus su	820	6	1.2	289	1	026566	026566 methanobact
748	6	1.2	256	2	095844	095844 staphylococ	821	6	1.2	289	2	045377	045377 caulobacter
749	6	1.2	256	5	007419	007419 typanosoma	822	6	1.2	289	2	067230	067230 aquifex aeo

823	6	1.2	289	11	062204	062204 mus musculu	896	6	1.2	320	10	Q41324	Q41324 stylosanthe
824	6	1.2	292	5	09XZLB	09XZLB drosophila	897	6	1.2	321	2	09WYL7	09WYL7 thermotoga
825	6	1.2	293	2	Q46564	Q46564 bacteroides	898	6	1.2	321	10	09XIV9	09XIV9 nicotiana t
826	6	1.2	293	2	Q52694	Q52694 roseburia c	899	6	1.2	321	11	054851	054851 mus musculu
827	6	1.2	293	2	09XBUD	09XBUD porphyromon	900	6	1.2	321	11	070610	070610 rattus norv
828	6	1.2	293	5	016423	016423 caenorhabd1	901	6	1.2	321	11	070610	070610 rattus norv
829	6	1.2	294	2	009257	009257 synechococc	902	6	1.2	322	2	Q46003	Q46003 campylobact
830	6	1.2	295	2	09WMRS	09WMRS pseudomonas	903	6	1.2	323	4	075028	075028 synechocyst
831	6	1.2	295	4	099322	099322 homo sapien	904	6	1.2	323	4	09Y698	09Y698 homo sapien
832	6	1.2	296	1	09YER8	09YER8 aeropyrum p	905	6	1.2	323	5	066835	066835 drosophila
833	6	1.2	296	2	053053	053053 listeria iv	906	6	1.2	323	11	088602	088602 mus musculu
834	6	1.2	296	5	017002	017002 caenorhabd1	907	6	1.2	324	1	029940	029940 archaeoglob
835	6	1.2	296	6	09ZT04	09ZT04 dicistonyx	908	6	1.2	324	10	Q41097	Q41097 populus tri
836	6	1.2	297	2	054383	054383 rickettsia	909	6	1.2	325	2	055392	055392 synechocyst
837	6	1.2	297	2	P71451	P71451 listeria mo	910	6	1.2	325	13	Q42169	Q42169 brachydanio
838	6	1.2	297	9	080085	080085 bacterioph	911	6	1.2	326	2	P76071	P76071 escherichia
839	6	1.2	297	10	Q41122	Q41122 phaeosul v	912	6	1.2	326	2	086322	086322 mycobacteri
840	6	1.2	298	1	059546	059546 pyrococcus	913	6	1.2	326	4	012251	012251 saccharomyc
841	6	1.2	298	2	P71702	P71702 mycobacteri	914	6	1.2	326	4	000272	000272 homo sapien
842	6	1.2	298	8	Q9XKX9	Q9XKX9 agrocyste ae	915	6	1.2	326	8	099019	099019 solanum dem
843	6	1.2	299	5	Q22856	Q22856 caenorhabd1	916	6	1.2	326	9	Q38213	Q38213 bacterioph
844	6	1.2	300	2	069782	069782 sinorhizobi	917	6	1.2	326	10	Q24392	Q24392 solanum tub
845	6	1.2	301	2	P94627	P94627 clostridium	918	6	1.2	326	10	055846	055846 methicoccc
846	6	1.2	301	2	Q44404	Q44404 agrobacteri	919	6	1.2	327	2	069447	069447 methicoccc
847	6	1.2	301	2	Q47396	Q47396 escherichia	920	6	1.2	329	5	018461	018461 haematobia
848	6	1.2	301	2	Q53756	Q53756 streptomyce	921	6	1.2	329	5	018461	018461 haematobia
849	6	1.2	301	8	09Y201	09Y201 ephydaria f	922	6	1.2	329	12	066100	066100 carnationa f
850	6	1.2	301	8	09Y207	09Y207 dicistonyx	923	6	1.2	329	12	09YMW5	09YMW5 lymantria d
851	6	1.2	302	5	018318	018318 caenorhabd1	924	6	1.2	330	1	058121	058121 pyrococcus
852	6	1.2	303	2	Q9ZG55	Q9ZG55 leptospira	925	6	1.2	330	2	Q9XSV7	Q9XSV7 streptomyce
853	6	1.2	303	8	Q9ZT05	Q9ZT05 dicistonyx	926	6	1.2	332	2	09X077	09X077 thermotoga
854	6	1.2	304	2	066991	066991 aquifex aeo	927	6	1.2	332	11	062627	062627 rattus norv
855	6	1.2	305	1	Q26838	Q26838 methanobact	928	6	1.2	333	2	09YWT8	09YWT8 thermotoga
856	6	1.2	305	2	067486	067486 aquifex aeo	929	6	1.2	334	2	Q25836	Q25836 heliobacte
857	6	1.2	305	2	007166	007166 mycobacteri	930	6	1.2	334	2	Q9ZT08	Q9ZT08 heliobacte
858	6	1.2	305	8	Q9ZT06	Q9ZT06 dicistonyx	931	6	1.2	334	3	014075	014075 schizosacch
859	6	1.2	305	13	Q9I346	Q9I346 xenopus lae	932	6	1.2	334	3	072733	072733 cowpox viru
860	6	1.2	306	2	Q47696	Q47696 escherichia	933	6	1.2	335	8	Q9ZV31	Q9ZV31 elephantiu
861	6	1.2	306	10	065623	065623 arabidopsis	934	6	1.2	336	1	Q30070	Q30070 archaeoglob
862	6	1.2	306	10	081845	081845 arabidopsis	935	6	1.2	336	2	Q49675	Q49675 mycobacteri
863	6	1.2	307	2	Q9ZT74	Q9ZT74 chlamydia p	936	6	1.2	336	2	P77745	P77745 mycobacteri
864	6	1.2	307	5	Q77394	Q77394 plasmodium	937	6	1.2	337	8	078674	078674 porphyra di
865	6	1.2	308	2	P72053	P72053 mycobacteri	938	6	1.2	337	4	000556	000556 homo sapien
866	6	1.2	308	2	067168	067168 aquifex aeo	939	6	1.2	337	5	077122	077122 laccosoma ch
867	6	1.2	308	5	Q45310	Q45310 caenorhabd1	940	6	1.2	338	2	Q45254	Q45254 bradyrhizob
868	6	1.2	309	5	Q23138	Q23138 caenorhabd1	941	6	1.2	338	2	P73887	P73887 synechocyst
869	6	1.2	309	10	P93550	P93550 spinacia ol	942	6	1.2	339	2	P94915	P94915 mycobacteri
870	6	1.2	310	2	052737	052737 streptococc	943	6	1.2	339	2	Q9XD85	Q9XD85 corynebacte
871	6	1.2	310	2	052738	052738 streptococc	944	6	1.2	340	2	09X7U9	09X7U9 streptomyce
872	6	1.2	310	2	052739	052739 streptococc	945	6	1.2	341	1	09YAE1	09YAE1 aeropyrum p
873	6	1.2	310	2	052740	052740 streptococc	946	6	1.2	341	5	095565	095565 homo sapien
874	6	1.2	310	2	052741	052741 streptococc	947	6	1.2	341	5	097103	097103 cithidia l
875	6	1.2	310	2	052743	052743 streptococc	948	6	1.2	341	5	097106	097106 trypanosoma
876	6	1.2	310	2	052745	052745 streptococc	949	6	1.2	341	5	097107	097107 trypanosoma
877	6	1.2	310	2	Q9ZBE7	Q9ZBE7 mycobacteri	950	6	1.2	342	1	058832	058832 pyrococcus
878	6	1.2	310	12	Q89181	Q89181 vaccinia vl	951	6	1.2	342	4	075796	075796 escherichia
879	6	1.2	312	2	Q25095	Q25095 heliobacte	952	6	1.2	342	4	075796	075796 homo sapien
880	6	1.2	312	4	095566	095566 homo sapien	953	6	1.2	342	8	098606	098606 porphyra le
881	6	1.2	312	4	016417	016417 caenorhabd1	954	6	1.2	342	12	Q55831	Q55831 stratford v
882	6	1.2	313	8	048005	048005 thiomomya ma	955	6	1.2	344	2	P94154	P94154 alcaligenes
883	6	1.2	314	2	Q33175	Q33175 mycobacteri	956	6	1.2	344	2	067238	067238 pseudomona
884	6	1.2	314	3	059909	059909 aspergillus	957	6	1.2	344	2	Q9Z9H2	Q9Z9H2 pseudomona
885	6	1.2	315	10	P93549	P93549 spinacia ol	958	6	1.2	344	5	017809	017809 caenorhabd1
886	6	1.2	315	10	Q24336	Q24336 raphanus sa	959	6	1.2	345	2	P75340	P75340 mycoplasma
887	6	1.2	316	5	018460	018460 haematobia	960	6	1.2	345	2	030883	030883 borrelia bu
888	6	1.2	316	5	018462	018462 haematobia	961	6	1.2	345	2	P73981	P73981 synechocyst
889	6	1.2	317	1	Q29954	Q29954 archaeoglob	962	6	1.2	345	6	Q27951	Q27951 bos taurus
890	6	1.2	317	2	066374	066374 acetobacter	963	6	1.2	346	8	078890	078890 anolis marc
891	6	1.2	317	2	066538	066538 aquifex aeo	964	6	1.2	346	2	Q25871	Q25871 heliobacte
892	6	1.2	318	1	057922	057922 pyrococcus	965	6	1.2	346	5	P91677	P91677 drosophila
893	6	1.2	319	2	006448	006448 serratia ma	966	6	1.2	347	8	019859	019859 porphyra am
894	6	1.2	320	2	068457	068457 acinetobact	967	6	1.2	347	8	P95603	P95603 alcaligenes
895	6	1.2	320	5	Q93616	Q93616 caenorhabd1	968	6	1.2	347	3	Q02428	Q02428 saccharomyc

```

969 6 1.2 347 3 012452
970 6 1.2 347 8 03165
971 6 1.2 347 8 03164
972 6 1.2 347 8 03087
973 6 1.2 347 8 03071
974 6 1.2 347 8 03263
975 6 1.2 347 8 03259
976 6 1.2 347 8 03253
977 6 1.2 347 8 03257
978 6 1.2 348 8 03255
979 6 1.2 348 5 02408
980 6 1.2 348 5 02408
981 6 1.2 348 5 02472
982 6 1.2 349 1 02472
983 6 1.2 349 4 02472
984 6 1.2 349 4 02472
985 6 1.2 349 10 02472
986 6 1.2 349 11 02472
987 6 1.2 349 12 02472
988 6 1.2 350 2 02472
989 6 1.2 350 2 02472
990 6 1.2 350 3 02472
991 6 1.2 350 5 02472
992 6 1.2 350 10 02472
993 6 1.2 350 12 02472
994 6 1.2 350 12 02472
995 6 1.2 350 12 02472
996 6 1.2 350 12 02472
997 6 1.2 350 12 02472
998 6 1.2 350 12 02472
999 6 1.2 350 12 02472
1000 6 1.2 350 12 02472

```

ALIGNMENTS

```

RESULT 1
ID 068697 PRELIMINARY; PRT: 326 AA.
AC 068697;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
DE V ANTIGEN.
GN ICRV.
OS Yersinia pestis.
OC Plasmid pCDL.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM;
RA HU P., ELLIOTT J., MCCREADY P., SKOMRONSKI E., GARNES J.,
RA KOHAYASHI A., CARBANO A.V., BRUBAKER R., GARCIA E.,
RA Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM;
RA MEDLINE: 9847122.
RA PERRY R.D., STRALEY S.C., FETHERSTON J.D., ROSE D.J., GREGOR J.,
RA "DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of
RT Yersinia pestis KIM5."
RL Infect. Immun. 66:4611-4623(1998).
DR EMBL: AF053946; AAC62574.1; -
DR EMBL: AF074612; AAC69799.1; -
KW Plasmid.
SQ SEQUENCE 326 AA; 37240 MW; 443BEDDC CRC32;

```

Query Match 46.6%; Score 243; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 7.1e-237;

```

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 ILKGGHYDNOQONIKRKEFLSSPTOMELRAFMVHPSLTADRIDDLKVIYDSM 338
DQ 84 ILKGGHYDNOQONIKRKEFLSSPTOMELRAFMVHPSLTADRIDDLKVIYDSM 143
QY 339 NHHGDANSKRLREELAEELAEIKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 398
DQ 144 NHHGDANSKRLREELAEELAEIKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 203
QY 399 EIRKASAEVILKMPOTTIOVDSSEKIKYISIKDFGSEKRRGALGNKNSYVYKDN 456
DQ 204 EIRKASAEVILKMPOTTIOVDSSEKIKYISIKDFGSEKRRGALGNKNSYVYKDN 263
QY 459 NELSHFATTCSDKSRPNDLVSOQKTQLOSDTSRNSAIEALNFIOKYDSVMORLDDT 518
DQ 264 NELSHFATTCSDKSRPNDLVSOQKTQLOSDTSRNSAIEALNFIOKYDSVMORLDDT 323
QY 519 SGK 521
DQ 324 SGK 326

```

```

RESULT 2
ID 056853 PRELIMINARY; PRT: 334 AA.
AC 056853;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE V ANTIGEN (FRAGMENT).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WA-314;
RA MEDLINE: 97162308.
RA ROEGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESEWANN J.,
RT "Passive immunity to infection with Yersinia spp. mediated by anti-
RT recombinant V antigen is dependent on polymorphism of V antigen."
RL Infect. Immun. 65:446-451(1997).
DR EMBL: X66796; CAA65588.1; -
FT NON_TER 334 334
SQ SEQUENCE 334 AA; 38241 MW; 71229360 CRC32;

```

```

Query Match 20.0%; Score 104; DB 2; Length 334;  
Best Local Similarity 100.0%; Pred. No. 1.5e-96;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 279 ILKGGHYDNOQONIKRKEFLSSPTOMELRAFMVHPSLTADRIDDLKVIYDSM 338
DQ 84 ILKGGHYDNOQONIKRKEFLSSPTOMELRAFMVHPSLTADRIDDLKVIYDSM 143
QY 339 NHHGDANSKRLREELAEELAEIKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 398
DQ 144 NHHGDANSKRLREELAEELAEIKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 203
RESULT 3
ID 056896 PRELIMINARY; PRT: 334 AA.
AC 056896;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE V ANTIGEN (FRAGMENT).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-NCTC;
 MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESESMANN J.;
 RT "Passive immunity to infection with *Yersinia* spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96797; CAA65589.1; -.
 FT NON_TER 334 334
 SO SEQUENCE 334 AA; 38227 MW; F23FED3 CRC32;

Query Match 20.0%; Score 104; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 1.5e-96;
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 ILKGHYDNLONGIRKRVKEFLESSPNTOWELRAFMVHFSLTADRIDDDILKVIYDSM 338
 |||
 DB 84 ILKGHYDNLONGIRKRVKEFLESSPNTOWELRAFMVHFSLTADRIDDDILKVIYDSM 143
 |||
 OY 339 NHHGDARSKLRELAELTAELKITYSVIOAEINKHLS 382
 |||
 DB 144 NHHGDARSKLRELAELTAELKITYSVIOAEINKHLS 187

RESULT 4
 ID 056846 PRELIMINARY; PRT: 324 AA.
 AC 056846;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE V ANTIGEN.
 OS *Yersinia enterocolitica*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Yersinia*.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y-108-P;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESESMANN J.;
 RT "Passive immunity to infection with *Yersinia* spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96799; CAA65591.1; -.
 FT NON_TER 324 324
 SO SEQUENCE 324 AA; 37194 MW; 5BD357FF CRC32;

Query Match 18.4%; Score 96; DB 2; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.8e-88;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 ILKGHYDNLONGIRKRVKEFLESSPNTOWELRAFMVHFSLTADRIDDDILKVIYDSM 338
 |||
 DB 84 ILKGHYDNLONGIRKRVKEFLESSPNTOWELRAFMVHFSLTADRIDDDILKVIYDSM 143
 |||
 OY 339 NHHGDARSKLRELAELTAELKITYSVIOAEINKHLS 374
 |||
 DB 144 NHHGDARSKLRELAELTAELKITYSVIOAEINKHLS 179

RESULT 5
 ID 056879 PRELIMINARY; PRT: 324 AA.
 AC 056879;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE V ANTIGEN.
 OS *Yersinia enterocolitica*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Yersinia*.
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESESMANN J.;
 RT "Passive immunity to infection with *Yersinia* spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96800; CAA65592.1; -.
 FT NON_TER 324 324
 SO SEQUENCE 324 AA; 37195 MW; E1CB300F CRC32;

Query Match 18.4%; Score 96; DB 2; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.8e-88;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 ILKGHYDNLONGIRKRVKEFLESSPNTOWELRAFMVHFSLTADRIDDDILKVIYDSM 338
 |||
 DB 84 ILKGHYDNLONGIRKRVKEFLESSPNTOWELRAFMVHFSLTADRIDDDILKVIYDSM 143
 |||
 OY 339 NHHGDARSKLRELAELTAELKITYSVIOAEINKHLS 374
 |||
 DB 144 NHHGDARSKLRELAELTAELKITYSVIOAEINKHLS 179

RESULT 6
 ID 056880 PRELIMINARY; PRT: 325 AA.
 AC 056880;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE V ANTIGEN (FRAGMENT).
 OS *Yersinia enterocolitica*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Yersinia*.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-8081;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESESMANN J.;
 RT "Passive immunity to infection with *Yersinia* spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96798; CAA65590.1; -.
 FT NON_TER 325 325
 SO SEQUENCE 325 AA; 37260 MW; D6E88446 CRC32;

Query Match 17.9%; Score 93; DB 2; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.9e-85;
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 ILKGHYDNLONGIRKRVKEFLESSPNTOWELRAFMVHFSLTADRIDDDILKVIYDSM 338
 |||
 DB 84 ILKGHYDNLONGIRKRVKEFLESSPNTOWELRAFMVHFSLTADRIDDDILKVIYDSM 143
 |||
 OY 339 NHHGDARSKLRELAELTAELKITYSVIOAEINK 371
 |||
 DB 144 NHHGDARSKLRELAELTAELKITYSVIOAEINK 176

RESULT 7
 ID 087495 PRELIMINARY; PRT: 324 AA.
 AC 087495;
 DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE ICRV PROTEIN.
 GN ICRV
 OS *Yersinia enterocolitica*.
 OG Plasmid pIV, and Plasmid pIVe227.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Yersinia*.
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-W22703;
RA IRIARTE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
RT "Complete sequence of the Yersinia enterocolitica pYV plasmid.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-W22703;
RA IRIARTE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
RT "Detailed genetic map of the pYV227 plasmid of Yersinia
   enterocolitica serotype O:9.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF102990; AAD16815.1;
KM plasmid.
SQ SEQUENCE 324 AA; 37296 MW; 7C9A6552F CRC32;

Query Match
Best Local Similarity 13.1%; Score 68; DB 2; Length 324;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 SYSYKDNNEISHFATTCSDKSRPLNDLVSOQTQLSDITSFNSAIFALNFIQKIDSV 510
DB 256 SYSYKDNNEISHFATTCSDKSRPLNDLVSOQTQLSDITSFNSAIFALNFIQKIDSV 315
QY 511 MQRLLDDT 518
DB 316 MQRLLDDT 323

RESULT 8
056882 ID 056882 PRELIMINARY; PRT; 324 AA.
AC 056882;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE Y ANTIGEN.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y-96-P.
RX MEDLINE; 97162308.
RA "ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESSEMAN J.;
   "Passive immunity to infection with Yersinia spp. mediated by anti-
   recombinant Y antigen is dependent on polymorphism of Y antigen.";
RL Infect. Immun. 65:446-451(1997).
DR EMBL: X96801; CAA65593.1;
SQ SEQUENCE 324 AA; 37162 MW; 56842DOA CRC32;

Query Match
Best Local Similarity 12.9%; Score 67; DB 2; Length 324;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHYDNLQNGIKRYKEFLSSPNTQWELRAFMAVHRSFLTRIDDDILKYIVDSM 338
DB 84 ILKGGHYDNLQNGIKRYKEFLSSPNTQWELRAFMAVHRSFLTRIDDDILKYIVDSM 143
QY 339 NHGDAR 345
DB 144 NHGDAR 150

RESULT 9
016256 ID 016256 PRELIMINARY; PRT; 168 AA.
AC 016256;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)

```

```

DE WT1-WILMS' TUMOR SUPPRESSOR PROTEIN.
GN WT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95166649.
RA HAMILTON T.B., BARIJLA K.C., ROMANIUK P.J.;
RT "High affinity binding sites for the Wilms' tumour suppressor protein
   WT1.";
RL Nucleic Acids Res. 23:277-284(1995).
DR EMBL: S75264; AAB33443.1;
DR HSSP; P08046; IALL.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR PFAM; PF00096; zf-C2H2; 4.
DR PRINTS; PR00048; ZINC_FINGER.
KM Zinc-finger: Metal-binding; DNA-binding.
SQ SEQUENCE 168 AA; 20165 MW; E86026CC CRC32;

Query Match
Best Local Similarity 3.3%; Score 17; DB 4; Length 168;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNHHHHHHHSSGHI 17
DB 1 MGNHHHHHHHSSGHI 17

RESULT 10
030527 ID 030527 PRELIMINARY; PRT; 294 AA.
AC 030527;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE PCRV.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
OC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-388;
RX MEDLINE; 98037517.
RA YAHN T.L., MENDE-MUELLER L.M., FRIESE M.B., FRANK D.W.;
RT "Identification of type III secreted products of the Pseudomonas
   aeruginosa exoenzyme S regulon.";
RL J. Bacteriol. 179:7165-7168(1997).
DR EMBL: AF010149; AAC45935.1;
SQ SEQUENCE 294 AA; 32283 MW; 0FAD2082 CRC32;

Query Match
Best Local Similarity 2.5%; Score 13; DB 2; Length 294;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 498 EALNFIQKIDSV 510
DB 274 EALNFIQKIDSV 286

RESULT 11
042186 ID 042186 PRELIMINARY; PRT; 516 AA.
AC 042186;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE FTZ-FI HOMOLOG.
GN ZFP1.
OS Brachydanio rerio (Zebrafish) (Zebra danio).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinidae; Rasbora; Danio.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 97322094.
 RA LUD D., LE DREAN Y., EKKER M., XIONG F., HEW C.L.;
 RT "Teleost F1 homolog and its splicing variant determine the
 RT expression of the salmon gonadotropin Ibeta subunit gene.";
 RL Mol. Endocrinol. 11:877-890(1997).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF014926; AAC60274.1; -.
 DR HSSP; P19793; 2ML.
 DR ZFIN; ZDB-GENE-990415-79; zff1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR PFAM; PF00104; hormone_rec; 1.
 DR PFAM; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STROIDFINGER.
 DR KEGG; K01492; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger.
 SQ SEQUENCE 516 AA; 58805 MW; 11AB2783 CRC32;

Query Match 2.3%; Score 12; DB 13; Length 516;
 Best Local Similarity 100.0%; Pred. No. 0.0017;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GHHHHHHHSS 14
 |||||
 DB 189 HHHHHHHHSS 200

RESULT 12
 ID 075263 PRELIMINARY; PRT; 230 AA.
 AC 075263;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE WIGSC.H.DJ0620P06.1 PROTEIN (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Euteleostei; Primates; Catarrhini; Hominiidae; Homo.
 RN [1]
 RA ANDREWS S., LACY M., DOBBELDE C.;
 RP Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA WATERSTON R.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; ACC05550; AAC33152.1; -.
 DR PFAM; PF00046; homeobox; 1.
 RT NON_TER 230 230
 FT SEQUENCE 230 AA; 25476 MW; 1A804F2B CRC32;

Query Match 2.1%; Score 11; DB 4; Length 230;
 Best Local Similarity 100.0%; Pred. No. 0.0081;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHH 12
 |||||
 DB 67 GHHHHHHH 77

RESULT 13
 ID 042980 PRELIMINARY; PRT; 420 AA.
 AC 042980;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE HYPOHETICAL 47.3 KD PROTEIN.
 GN SPBC17D11.01 OR SPBC20F10.11.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972H-;
 RA BECK A., REINHARDT R., WOOD V., RAJANDREAM M.A., BARRELL B.G.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-281 FROM N.A.
 RC STRAIN-972H-;
 RA LYNE M., MCDONALD R., RAJANDREAM M.A., BARRELL B.G., BECK A.,
 RA REINHARDT R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL031739; CAA21072.1; -.
 DR EMBL; AL021747; CAA16851.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 420 AA; 47257 MW; 35DC8FC CRC32;

Query Match 2.1%; Score 11; DB 3; Length 420;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHH 12
 |||||
 DB 298 GHHHHHHH 308

RESULT 14
 ID P78807 PRELIMINARY; PRT; 427 AA.
 AC P78807;
 DT 01-MAY-1997 (TReMBLrel. 03, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE FISSION YEAST.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PR745;
 RX MEDLINE; 98162722.
 RA YOSHIOKA S., KATO K., NAKAI K., OKAYAMA H., NOJIMA H.;
 RT "Identification of open reading frames in Schizosaccharomyces pombe
 RT cDNA.";
 RL DNA Res. 4:363-369(1997).
 DR EMBL; D89156; BAA13818.1; -.
 FT SEQUENCE 427 AA; 48056 MW; 47E3B3FE CRC32;

Query Match 2.1%; Score 11; DB 3; Length 427;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHH 12
 |||||
 DB 298 GHHHHHHH 308

RESULT 15
 ID 09YGC6 PRELIMINARY; PRT; 467 AA.
 AC 09YGC6;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)

OM of: US-08-699-716a-2 to: GenEmbl.* out-format : pfs

Date: Aug 22, 2000 5:22 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=frame+pn.model -DEV=xlp
-O/cg22.1/USPTO.spool/US08699716/runat.31072000.084837.24695/app-query.fasta.1.587
-B-genEmbl -QEMT-fastap -SUFFIX=rg -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPT=0.000 -LOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -DELDP=6.000
-DELKXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELDP=6.000
-DELKXT=7.000 -START=1 -MATRIX=bloms62 -TRANS=human6.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTEMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=100000 -USER=US08699716 -CGENL_1=3898 -NCPD=6 -ICPD=3
-LONGLOG -NO_XLXPY -WAIT -THREADS=1

Search information block:

Query: US-08-699-716a-2

Query length: 521

Database: GenEmbl.*

Database sequences: 972840

Database length: 892348106

Search time (sec): 1190.920000

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat:A56814	2482.50	2626.99	3.1e-138	1530	A56814 Sequence 22 from Patent
gb_pat:A56808	2387.50	2526.98	1.2e-132	1462	A56808 Sequence 16 from Patent
gb_pat:YEPBCR	1646.00	1740.05	7.9e-89	2100	M2403 Yersinia pestis lcrG, l
gb_pat:YPCD1	1646.00	1709.08	4.2e-87	70305	A1117189 Yersinia pestis plas
gb_pat:AF053946	1646.00	1709.05	4.2e-87	70504	AF053946 Yersinia pestis plas
gb_pat:AF074612	1646.00	1709.05	4.2e-87	70559	AF074612 Yersinia pestis plas
gb_pat:A56411	1641.50	1741.72	6.4e-89	1014	A56411 Sequence 1 from Patent
gb_pat:A56793	1641.50	1741.19	6.4e-89	1014	A56793 Sequence 1 from Patent
gb_pat:A46413	1641.00	1741.19	6.9e-89	1014	A46413 Sequence 3 from Patent
gb_pat:A56795	1641.00	1741.19	6.9e-89	1014	A56795 Sequence 3 from Patent
gb_pat:AF167309	1631.00	1730.96	2.5e-88	975	AF167309 Yersinia pestis strat
gb_pat:AF167310	1628.00	1727.74	3.9e-88	981	AF167310 Yersinia pestis strat
gb_pat:YEPYVNT	1590.00	1687.57	6.7e-86	981	X96802 Y. pseudotuberculosis V a
gb_pat:YEPBCRVHP	1590.00	1680.45	1.7e-85	2201	M57893 Yersinia pseudotubercul
gb_pat:AF102990	1537.00	1615.05	7.3e-82	69673	AF102990 Yersinia enterocolit
gb_pat:YEP314VANT	1552.50	1647.75	1.1e-83	1002	X96796 Y. enterocolitica V anti
gb_pat:YENCYVANT	1549.50	1644.58	1.7e-83	1002	X96797 Y. enterocolitica V anti
gb_pat:YEP808VANT	1537.00	1631.61	8.7e-83	975	X96798 Y. enterocolitica V anti
gb_pat:YEP96PVANT	1532.00	1626.32	1.7e-82	975	X96801 Y. enterocolitica V anti
gb_pat:YEP108VANT	1531.00	1625.27	2.0e-82	975	X96799 Y. enterocolitica V anti
gb_pat:YEP527VANT	1527.00	1621.04	3.4e-82	975	X96800 Y. enterocolitica V anti
gb_pat:A56812	852.00	892.69	9.7e-43	5383	A56812 Sequence 20 from Patent
gb_pat:YPCAF	852.00	892.52	1.3e-41	5383	A56196 Y. pestis genes catI, ca
gb_pat:YPPMT1	852.00	867.09	3.3e-40	96210	A1117211 Yersinia pestis plas
gb_pat:AF053947	852.00	866.66	3.5e-40	100984	AF053947 Yersinia pestis plas
gb_pat:AF074611	852.00	866.66	3.5e-40	100980	AF074611 Yersinia pestis plas
gb_pat:AF010149	490.00	512.15	2.0e-20	4168	AF010149 Pseudomonas aeruginos
gb_pat:SC8727	352.00	391.47	1.0e-13	240	S38727 lcrGVH Operon: lcrV-V-ant
gb_pat:YEPFARPCB	164.50	169.03	0.2551	3755	M88098 Plasmidium vlyax relict
gb_pat:155033	164.50	169.01	0.2557	3763	I155033 Sequence 1 from Patent
gb_pat:AF081364	164.00	168.84	0.0382	660	AF081364 Synthetic construct CD
gb_pat:AR000030	154.00	166.62	0.3474	1402	AR000030 Sequence 23 from pate
gb_pat:SPCC417	152.50	133.40	19.06	40383	AL0035076 S. pombe chromosome I
gb_pat:SCYLO87C	150.50	133.06	2.25	4802	Z74829 S. cerevisiae chromosome
gb_pat:SCCHXVORF	150.50	133.52	24.25	39307	X31321 S. cerevisiae DNA for C
gb_pat:AC016121	150.50	129.33	131.94	175733	AC016121 Homo sapiens clone
gb_pat:AC026332	150.50	129.27	150.77	197726	AC026332 Homo sapiens chromo
gb_ov:CHRC7B	150.00	155.92	1.37	2920	D42150 Chicken mRNA for cadher
gb_ov:AF014012	149.00	150.79	2.65	4638	AF014012 Gallus gallus reelin
gb_pat:AF001153	149.00	139.29	11.57	17070	AF001153 Borrelia burgdorferi
gb_pat:AP000466	149.00	122.12	104.61	119580	AP000466 Homo sapiens genom

gb_in2:AF079445 + 148.50 141.39 8.83 12868 I AF079445 Dictyostellium disc
gb_in2:AF057019 + 147.00 143.25 6.96 8574 I AF057019 Dictyostellium disc
gb_htg15:AC015834 - 145.50 114.92 263.61 177968 I AC015834 Homo sapiens clo
gb_p11:SCYLO90W + 145.00 149.74 3.03 3234 I Z74832 S. cerevisiae chromos

seq.name: gb_pat:A56814

seq.documentatn_block:

LOCUS A56814 1530 bp DNA PAT 03-MAR-1998

DEFINITION Sequence 22 from Patent WO9628551.

ACCESSION A56814

VERSION A56814.1 GI:3112827

KEYWORDS Yersinia pestis.

SOURCE Yersinia pestis

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE 1 (bases 1 to 1530)

AUTHORS Tithball, R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C.,

Bennett and Alice, M.

TITLE VACCINES FOR PLAGUE

JOURNAL Patent: WO 9628551-A 22 19-SEP-1996;

COMMENT SECR DEFENCE (GB)

Other publication ZA 9602036 960716

Location/Au 4951196 961002.

FEATURES

source

CDS

/db_xref="taxon:632"

/note="unnamed protein product"

/codon_start=1

/transl_table=1

/protein_id="CA03423.1"

/db_xref="GI:3112828"

/translation="MKRISSVAIALEGTIAANADLASTATATVAPRITLY

KEAPPTIDKNDNGIDIELVGLTGLGVTGTSVNTDAAGDMVLTFTSDGN

HOTTATVIGKSDPDSPKNGENGVGVATGSDPFVSIKSGKLAAGKT

DATVTVSNQSGTEGRIRAYEONPHIEDLEKRYEOLTGSSVLEELVOLKRD

IDSTVDPKSEVRANVITDDILKRIAYLPEDALIKGSHVNDOLNGIRV

KEPTESSPTQWELRPFVAMFSLADRIDDLKVIYDSNMHHDASKLEELAE

LVAELIVSVQAEIKHLSSGTHINRHSINLMDKNISYGTDEEIPASAKYLE

KMPTQIVDSGKTKIVSKIDFGSENKTKLGNLKNISYKDNNEISHPATYICSD

KSPINDLVISOKTQSDITSFNSIALNRFIOYDVMRLDDTSGK"

BASE COUNT

500 a 296 c 311 g 423 t

ORIGIN

alignment_scores:

Quality: 2482.50

Ratio: 4.995

Percent Similarity: 99.004

Percent Identity: 98.606

alignment_block:

US-08-699-716A-2 ..

Align seg 1/1 to: A56814 from: 1 to: 1530

23 HismetLysLysLysSerValIleAlaIleAlaLeuphegLYthrl 39

10 TATATGAATAAATAGTTCCTATGCCATGTCATATTGGAACAT 59

39 eAlaIhrAlaAsnAlaAlaAspLeuThrAlaSerThrThrAlaTrAlat 56

60 TGCAACTCTCTATGCGGAGATTACTGCAAGCCACTCCAAAGCGGAA 109

56 hLeuValGlpProAlaArgIleThrLeuThrTyLysGlgLYAlaPro 72

110 CTTCTGTGACACCGCCGATCTTCTATATATAGGAAGGCCCTCCA 159

73 IlerhrlIleMetAspAsnGlyAsnIleAspThrIleLeuLeuValGlyTh 89

160 ATTCAATATATGCAATGAAACATGATACAGATATTGTTGTTGAC 209

89 rLeuThrLeuGlyGlyTyrIleThrGlyThrThrSerThrSerValAsp 106
 110 GCTTACTTGGCGGCTATAAAGAGAACACCTAGACATCTGTACT 239
 106 hethrAspAlaAlaGlyAspProMetTyrLeuThrPheThrSerGlnAsp 122
 260 TTTACAGATGCCGGGGTGATCCCATGCTTACATTACTTCTCAGCAT 309
 123 GlyAsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArgAs 139
 310 GGAATAATACCAACCATCTACTACAAAAGTATGGCAAGATCTTAGAGA 359
 139 PheAspPheLeuSerProLysValAsnGlyLysLeuValGlyAspAsp 156
 360 TTTTATATCTCTCTAAGGTAAACGGTGAACCTTGGGGGATGACG 409
 156 aValLeuAlaThrGlySerGlnAspPheValArgSerIleGlySer 172
 410 TCGTCTTGGCTACGGGACCCAGGATTTCTTGTCCCTCAATTGGTTCC 459
 173 LysGlyGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 189
 460 AAGGGGGTAACTGCGAGGAAATACATGATGCTGAACCGTAAAC 509
 189 rValSerAsnGln.....GluPheMetIleArgAlaTyrGluGln 203
 510 CGTATCTAACCAAGATCCATCAAGTCTGATTAGAGCTTACCAACAA 559
 203 snProGlnHisPheIleGlyAspLeuGlyLysValArgValGluGlnLeu 219
 560 ACCCACACACTTTTATTGAGGATCTAGAAAAGTTAGGTGGAAACAATT 609
 220 ThrGlyHisGlySerSerValLeuGluGluLeuValGluLeuValLysAs 236
 610 ACTGTCATGCTTCTTCAAGTTTGAAGAATGGTTCAGTTAGTCAAGA 659
 236 PheAsnIleLeuAspLeuSerIleLysTyrAspProArgLysAspSerLys 253
 660 TAAATATATAGATATTTCCATTAATATGATCCCAAGAAAGATTCGAGG 709
 253 alphaLeuAsnArgValIleThrAspAspIleGluLeuLeuLysLysIle 269
 710 TTTTTCGCAATAGATATATCTGATGATATCCAAATGCTCAAGAAATC 759
 270 LeuAlaTyrPheLeuProGluAspThrIleLeuLysGlyLysIleThrAs 286
 760 CTAGCTTATTTCTACCCGAGATGCCATCTTAAAGCGGTCAATTATGA 809
 286 PheGlnLeuGlnAsnGlyLysLysArgValLysGluPheLeuGlnLys 303
 810 CAACCAACTGCAAAATGCGATCAAGCAGATMAAGAGTCTCTGAAACAT 859
 303 serProAsnThrGlnTyrGluLeuArgLysPheMetAlaValMetHisPhe 319
 860 CGCGGATATACAAATGGAATTCGGGCTTCATGCAAGTAATGCAATTC 909
 320 SerLeuThrAlaAspArgLysAspAspIleLeuLysValIleValAs 336
 910 TCTTTAACCCCGCATCTGATGATGATATTTTGAAGTATGTTTGA 959
 336 PseMetAsnHisHisGlyAspAlaArgSerLysLeuArgGluGluLeu 353
 960 TTTAATGAATCATCATGCTGATGCGCTACCAAGTTGCGGAAGAAATTAG 1009
 353 IacLeuLeuThrAlaGluLeuLysIleTyrSerValIleGlnAlaGluIle 369
 1010 CTAGACTTACCGCGCAATTAAAGATTTATTCAGTTATTCACCGCAAAAT 1059
 370 AsnLysHisLeuSerSerSerGlyThrIleAsnIleHisAspLysSerIle 386
 1060 AATAACATCTGTCTAGTAGTGCACCATTAATATCATGATTAATCAT 1109

386 eAsnLeuMetAspLysAsnLeuTyrGlyTyrThrAspGluGluIlePhe 403
 1110 TAACTCATGATTAATAAATTTATATGTTATACAGATGAAGAAATTTTA 1159
 403 ysaIleSerAlaGluTyrLysIleLeuGluLysMetProGlnThrThrIle 419
 1160 AAGCCAGCGGAGAGTCAAAAATTCGAGAAAATGCTCAAAACCACTT 1209
 420 GlnValAspGlySerGluLysLysIleValSerIleLysAspPheLeu 436
 1210 CAGGTGATGGAGCGAGAAAAAATAGTCTGATTAAGAGACTTCTTGG 1259
 436 ySerGluAsnLysArgThrGlyAlaLeuGlyAsnLeuLysAsnSerTyrS 453
 1260 AAGTGAATAAAGAACCGGGCGTGGGTATTCGAAAAACTCATACT 1309
 453 eTyrAsnLysAspAsnAsnGluLeuSerHisPheAlaThrThrCysSer 469
 1310 CTTATATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1359
 470 AspLysSerArgProLeuAsnAspLeuValSerGlnLysThrThrGln 486
 1360 GATTAAGTCCAGCGGCTCAACGACTTGTGTTAGCAAAAAACAACACTG 1409
 486 uSerAspIleThrSerArgPheAsnSerAlaIleGluAlaLeuAsnArgP 503
 1410 GTCTGATATTACATCACTTATTAATCAGCTATGTAAGCACTGAACCTT 1459
 503 helLeuGlnLysTyrAspSerValMetGlnArgPheLeuAspAspThrSer 519
 1460 TCATTTAGAAATATGATTCAGTATGCAACGTCCTTAATGACAGCTCT 1509
 520 GlyLys 521
 1510 GGTAAA 1515

seq_name: gb_pat:A56808

seq_documentation_block: 1462 bp DNA

LOCUS A56808

DEFINITION Sequence 16 from Patent WO9628551.

ACCESSION A56808

VERSION A56808.1 GI:3712821

KEYWORDS

SOURCE

ORGANISM

Yersinia pestis.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.

REFERENCE 1 (bases 1 to 1462)

AUTHORS Tlball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,

Bennett and Alice,M.

TITLE VACCINES FOR PLAGUE

JOURNAL Patent: WO 9628551-A 16 19-SEP-1996;

SECUR DEFENCE (GB)

Other publication ZA 9602036 960716

Other publication AU 4951196 961002.

location/Qualifiers

1..1462

/organism="Yersinia pestis"

/db_xref="taxon:632"

<8..1450

/note="unnamed protein product"

/codon_start=1

/transl_table=1

/protein_id="CA03421.1"

/db_xref="GI:3712822"

/translation="ADTPASTTATATATVEPARITLTYKEGAPITMDNGNIDELLVG

TLTIGGKRTGTSVNFETDAAGDPMILFTSODGNHOFRTKVGDSRDFDISPV

NGENLVGDVVLATGSDFFVRSIGSGKLAAGKYDAVTVVSNOSGSEGRIRAVE

QNPHEFLEDEKVEVLEDTGHSVLELVOLVDKNDISIKYDPRKDSVFANRVT

TDDELKRTIAYFLPDAITLKGHYNOICONGIKRVEKFELESPTOMELRAFMAY

HFSYTAIRIDDDILKIVYDSNNHGDARSKIRELALETAEKITYSTIQAEINKHLS

SGTINIDKSLINLMDKNLYGTTDEIFKASAEYKILKEKPOTTIQVDSSEKKIVSIND

BASE COUNT	476 a	286 c	300 g	400 t
ORIGIN	FEGSENRGTGALGNLKNSYSYXNDNNNSHEATTCSDKSRRLNDLVSOQKTQGLSDITS RFNSAITALNFIQIKYDSVMORLLDDTSGK"			

```

alignment_scores:
  Quality: 2387.50      Length: 4882
  Ratio: 5.005         Gaps: 1
Percent Similarity: 98.963  Percent Identity: 98.340

```

```
alignment_block:
```

Align seg 1/1 to: A56808 from: 1 to: 1462

43	ASPAALALASPLENTHRALSERTHRLATRLATRLALVAILGI	59
2	ACGTGGCAGATTTAACGTCAAGCACCACTGCACGGCACTCTGTGA	51
59	UPFOALAAAGLIERHLEUTHTYTLYSGUGLYALABROLLEHRLIEM	76
52	ACACAGCCCCACACTCTTACATATAAAGGAGCGCTCAATTACAAATTA	101
76	ETSPASNDLYASNLIASPHRLIULEULEVALGLYTHLEUTHLEU	92
102	TGACACATGGAACATCGATACAGAAATTACTGTGGTACCTTACCTT	151
93	GLYGLTYTLYSGTHGLYTHRHRSERTHSERVALASPHETRRASPAI	109
152	GGCGGCTATATAAAGAGAACCACTAGACACACTGTGTACTTACAGATGC	201
109	AALAGIYASPPROMETTYTLEUTHPHETHSERGLNASPLYNASNH	128
202	CGGGGTGATCCCATGTACTTAACTTACCTTCAGGATGGAATTAACC	251
126	ISGLNPHETHRLHYLSVALILEGLYASPSERASPPHASPILE	142
252	ACCAATTCCTACAAAAAGATGGCCAAAGATTCAGAGATTTTGATATC	301
143	SERPROLYSVALASNDLYGLUNSDLEUVALGIYASPAVALVALLUAI	159
302	TCCTCTTAAGGTAAACGCTGAGAACTTGTGGGGATGACGTCGCTTGGC	351
159	ATHRGYSERGLNASPPHPEVALATRSERILEGLYETLYSGELYLYL	178
352	TACGGGCACCCAGGATTTCTTGTCGCTCAATGGTGTCCAAAGCGGTA	401
176	YSLEUALAAGLYSTYTHRRASPAALVALTHVALTHVALSERASN	192
402	AACCTTCACAGGTAAATCACTGATGCTGTACCCGTAAACCTTATCTAAC	451
193	GLN.....GLUPHEKELIETALRYALATYRGLUGLINAASPROGINH1	206
452	CAAGGATTCATCGAAGAGGTGTTATAGAGCTTACGAAACCAACCAACA	501
206	SPHELIEGLIUSPBLEUGLULYSVALATRYALGLUGLINDLEUTHRLYHSG	222
502	TTTTATTGAGACTTATGAAAAAGTTAGGTGGAAACACTCTGTGTATG	551
223	LYSESEERVALLEUGLUGLULEUVALGLINLEUVALIYASPLYSASPILE	238
552	GTTCTTCAGTTTATAGAGAAATGGTTCAGTTAGTCAAGATTAATAATA	601
240	ASPIIISERILEUSTYRASPPROAGLYSASPSERGLUALPHEALIAS	256
602	GAAATTTCAATTAATAATATATCCAGAAAAAGATTGGGAGTTTGGCAA	651
256	NATRYVALIETHRASPSPIIGLULEULEUYSYSLILEULALATYRP	273
652	TAAAGTATATTACGATGATGATGCAATGCTCTCAAGAAAATCTTACCTTATTT	701
273	HELEUPROGLUASPHRLILEULEUYSGLYGLYHISTYRASPSASNDLEU	289

seq_name: gb_bai:YEPLC1

LOCUS YEPLC1 2100 bp DNA BCT 26-APR-1993
DEFINITION Yersinia pestis lcrF, lcrV, and lcrH genes, complete cds.
ACCESSION M26405.1 GI:155448
VERSION 1
KEYWORDS lcrF protein; lcrH protein; lcrV protein; V antigen.
SOURCE 75kb virulence plasmid.
ORGANISM Yersinia pestis
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersiniinae.

REFERENCE
1 (bases 1 to 2100)
Price,S.B., Leung,K.Y., Barve,S.S. and Straley,S.C.
AUTHORS

TITLE Molecular analysis of *lcrGH*, the *V* antigen operon of *Yersinia*
 JOURNAL *J. Bacteriol.* 171, 5646-5653 (1989)
 MEDLINE 90008086
 COMMENT Computer readable copy of sequence [*J. Bacteriol.* (1989) In press]
 kindly submitted by Price, S.
 07-AUG-1989.

FEATURES
 source location/Qualifiers
 1..2100

/organism="Yersinia pestis"
 /strain="KIM5"

/db_xref="taxon:632"

-35_signal
 -10_signal

147..152

174..179

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

192..479

alignment_scores:

Quality: 1646.00

Ratio: 5.049

Percent Similarity: 100.000

Percent Identity: 99.693

Length: 326

Gaps: 0

Percent Identity: 99.693

Percent Identity: 99.693

Percent Identity: 99.693

Percent Identity: 99.693

Percent Identity: 99.693

BASE COUNT 688 a 423 c 461 g 528 t

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

|||||

531 AAAGTTAGGCTGCAACACTTACTGTCATGTTTTCAGTTTGAAG 580

229 ludeuvalglnleuvallyasplyasnlleasplleserileystyr 245

581 AATTGCTGCTAGTCAAGATATAAATATATATATTCATTATAT 630

246 ASPProarglyaspserserlyvalphealaasnargvallethraps 262

631 GATCCAGAAAGATTCGAGGTTTTCGCAATAGATATCTGATGA 680

262 plleuvalleuvallyasplleuvaltyrpheluproglyaspthr 279

681 TATGATTCGCTCAGAAATCTTACTTATTTTCCAGGAGATGCA 730

279 leuvalsglyglyh1styraspasnleuvalasnnglyileysarg 295

731 TTCTTAAGCGGCTCATATGACACCACTGCAAAATGCAATCAAGCA 780

296 vallysgluupheleugluserproasnthrtrpdludeuargal 312

781 GTAAAGAGTTCCTGATCATCGCGAATACACATGGGAATGCGGCG 830

312 aphelealavalmethispheserleuthralaasparglyleaspa 329

831 GTTCATGGCAGTATGATTCCTTTTAAACCGCGATCGATCGATGATG 880

329 spleuvallylevalaspsersermetasnshisglyaspararg 345

881 ATATTTTGAAGTATGATTCGATTCATATGATATGATGATGATGAT 930

346 Serlyseuvalrgluuvalleuvalgluvalleuvalleuvalleu 362

931 AGCAAGTGGCGTGAATAATAGCTGAGCTTACCGCGAATTAAGATTA 980

362 rservalileglalagluileasnlshisleserleuvalleuval 379

981 TTGAGTATTCAGCCGAAATTAATAGCATCTGTAGTAGGACCA 1030

379 leasnlehisasplyserleasnleuvalasplyasnleuvalrgly 395

1031 TAATATCCATGATTAATCATATCATGATTAATTAATTAATGAT 1080

396 TyrtThrAspGluIuIephelyasalaSerAlaGluTyrIleLeu 412

1081 TATACAGATGAAGATTTTAAAGCAGGACGAGATACAAATTCCTCG 1130

412 uysmetproglthrthrleuvalasplysergluylslyleuval 429

1131 GAAATATGCTCAACCACTTACGATGAGTGGACGAGAAAAATAG 1180

429 alserileuasppheleuglysergluasnlyasrgrhlyalaleu 445

1181 TCTCGATTAAGGACTTCTTGAGAGTGAATTAAGAACCGGGGCGTGG 1230

446 Glyasnleuvalasnseryserlyasnllyasasnleuvalleu 462

1231 GGTATCTGAAAAAATCATCTCTTATTAATTAAGATTAATTAATTA 1280

462 rHisphealaThrThrCysSeraspIysSerArgProLeuAsnAspLeu 479

1281 TCACCTTGGCCACCGCTGCTCGATTAAGTCCAGCGCTCAACGACTTGG 1330

479 alsergluylsThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495

1331 TTAGCCAAAAAACAATCACTGCTCTGATATTCATCAGCTTTTAATTA 1380

496 AlaIleGluAlaLeuAsnArgPheIleGlnIleTyrAspSerValMetG 512

1381 GCTATTGAAGCACTGAACCGTTCATTCAGAAATATGATTCAGTATGA 1430

512 nArgLeuLeuAspAspThrSerGlyLys 521

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

1431 ACGTCTGCTAGATGACACGCTGTAA 1458

seq_name: gp_baz:YPCD1

seq_documentation_block:

LOCUS YPCD1 70305 bp DNA BCT 22-MAR-2000

DEFINITION Yersinia pestis plasmid pCD1.

ACCESSION AL117189

VERSION AL117189.1 GI:5832423

KEYWORDS chaparrone; cytotoxic effector; IS100; IS1616; IS1617; lcr; low-calcium response; syc; targeted effector; transposase; type III secretion; Y antitoxin; virulence; Ylp; Yop; ysc.

SOURCE Yersinia pestis

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.

REFERENCE 1 (bases 1 to 70305)

AUTHORS Karlyshev,A.V. and Wren,B.W.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 70305)

AUTHORS Baker,S.G. and Mungall,K.

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 70305)

AUTHORS James,K.D., Parkhill,J., Barrell,B.G. and Randalream,M.A.

JOURNAL Submitted (03-SEP-1999) Yersinia pestis sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA

REFERENCE 4 (bases 1 to 70305)

AUTHORS E-mail: barrell@sanger.ac.uk DNA supplied by Dr. Andrey Karlyshev and Prof. Brendan Wren, (3). Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT

REFERENCE 5 (bases 1 to 70305)

AUTHORS Notes:

Yersinia pestis sequencing at The Sanger Centre is funded by Beowulf Genomics

Details of Y. pestis sequencing at the Sanger Centre are available on the World Wide Web.

(URL: http://www.sanger.ac.uk/Projects/Y_pestis/)

CDS are numbered using the following system eg YPCP1.01c. YP (Y. pestis), PCP1 (plasmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.

The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database.

The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon.

CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or att) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

Location/Qualifiers

1. 70305

/organism="Yersinia pestis"

/plasmid="pCD1"

/strain="CO-92 B1ovar Orientalis"

/db_xref="taxon:632"

1. 1956

/note="IS100 element"

88. 1110

/gene="YPCD1.01"

88. 1110

/gene="YPCD1.01"

/note="YPCD1.01, probable transposase, len: 340 aa; putative insertion sequence IS100, identical to corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946, AF074612) (340 aa), fasta scores: opt: 2328 z-score: 3808.9 E(): 0. 100.0% identity in 340 aa overlap. Similar to many others e.g. TRAO_ECOLI (EMBL:X14793), IS1b, E.coli transposase for insertion sequence element IS21 (390 aa) (33.1% identity in 329 aa overlap). Contains P1am match to entry PF00239 recombinase, Site-specific recombinases,

score 25.70, E-value 4.8e-06. Contains probable helix-turn-helix motif at aa 19-40 (Score 2045, +6.15 SD)''

/transl_table=11

/label="YPCD1.01"

/product="putative transposase"

/protein_id="CAB54878.1"

/db_xref="GI:5832424"

/translation="ATFETVWEKILKKGSSRAIRBELGINSRTYKRYLAKSEP EKYTRPAPVASLDEYRDYINRIRADAPKYPKIPAVINAREIDQYRGKMTLRFIR SISVQEDPAPVAFETEPGRQVQWGTNRKSPHVAVLGTSRLYIEFTNMR YDLETCRHNAPREFGVPREYLDNMATVYLRDAPYQGTGRFSPIMQEKGFSPICRPFRAQTCKVERMYOYRNSFYIPMLTRLRPMGITVDVETVARNRNGHLMIDVA NQRKHETIOARPCDRMLEEOGSMALPPEKKEYDVHLDENLVNFKHPLHPLSYDS FCRGVA"

112. 195

/gene="YPCD1.01"

/note="P1am match to entry PF00239 recombinase, site-specific recombinases, score 25.70, E-value 4.8e-06"

1110. 1889

/gene="YPCD1.02"

1110. 1889

/gene="YPCD1.02"

/note="YPCD1.02, probable transposase, len: 259 aa; putative insertion sequence IS100, identical to corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946, AF074612) (259 aa), fasta scores: opt: 1658 z-score: 2693.4 E(): 0. 100.0% identity in 259 aa overlap. Similar to many others e.g. IS1b_ECOLI (EMBL:X14793), IS1b, E.coli transposase for insertion sequence element IS21 (265 aa) (47.4% identity in 249 aa overlap). Contains PS00017 ATP/GTP-binding site motif A (P-loop)''

/codon_start=1

/transl_table=11

/label="YPCD1.02"

/product="putative ATP-binding protein"

/protein_id="CAB54879.1"

/db_xref="GI:5832425"

/translation="AMELOHQMLMLAGQLQLESLSAARPAISQAVQDQSYMPLE HILHEKRLARRKQATYTRMAFPVAVTFEYTDFTFTNGAPKQLOSLRSFTERN ENIVLVGSGVGRKTLAIAAGYEAIVAGIKVETFDLADLLSTAGRGKRTYLOR GVMAPRLILIDIGILPFSOEAKLFFVIARKEYKSAMILTSNLPFGQMDTFAGDA AITSAMDLRIHSHVVOIKGSEYRLRQKRKGVIAEANPE"

1434. 1457

/gene="YPCD1.02"

/note="PS00017 ATP/GTP-binding site motif A (P-loop)''

1956. 46271

/note="Identical to Y.pestis KIM5 plasmid pCD1 (EMBL:AF053946) from 46489 to 1955, and to Y.pestis KIM5 plasmid pCD1 (EMBL:AF074612) from 59097 to 14563, except where noted"

complement(1956..2204)

/gene="YPCD1.03c"

complement(1956..2204)

/Partial

/gene="YPCD1.03c"

/note="YPCD1.03c, probable transposase remnant, len: 83 aa; similar to several e.g. N-terminus of TRAO_ECOLI (EMBL:X14793), IS1a, E.coli transposase for insertion sequence element IS21 (390 aa), similarity is interrupted by the adjacent IS100 element. The remainder of this CDS is in YPCD1.97c. This region is also similar to TR-068707 (EMBL:AF053946, AF074612) Y.pestis KIM5 putative transposase in pCD1 (390 aa). Contains probable helix-turn-helix motif at aa 19-40 (Score 1833, +5.43 SD)''

/codon_start=1

/transl_table=11

/label="YPCD1.03c"

/protein_id="CAB54880.1"

/db_xref="GI:5832426"

/translation="MISREDFYMIKMRQOAGYIIDIAOVGCSERTVRRYLKPEPP AKTRHKWKLKPFMDYIDIMRLAEVNNSEVILAIKAM"

2304. 2948

gene

```

CDS
/gene="YPCD1.04"
2304..2948
/partial
/gene="YPCD1.04"
/feature="YPCD1.04, possible transposase remnant, len: 215
aa: similar to many e.g. TR:Q4612 (EMBL:X78052)
Enterobacter agglomerans IS 1222 ORF8 (276 aa). Truncated
at N-terminus"
/codon_start=1
/transl_table=11
/label="YPCD1.04"
/protein_id="CAB54881.1"
/db_xref="GI:5832427"
/translation="HRSGLVHNKRVYRLVHLSGLGVKRRRRKGLATERLPILPRA
PNLTWMDVFDALATGRRKICVVDYRECLTVAFGISGVQVRIIDSLER
GYPATIRIDQPEFTCRALDQAFHGVHLLIQPKETQNGFIIESFNGRDECLNE
HMFSDVSHARKTISEMRDYNCRPHSLNQTSEFAAARKNSDESGDITK"
3004..2982
/misc_feature
/feature="similar to E.agglomerans IS 1222 (EMBL:X78052) at
DNA level"
complement(3014..3406)
/gene="YPCD1.05c"
complement(3014..3406)
/gene="YPCD1.05c"
/feature="YPCD1.05c, syce, yerh, yope targeting protein, len:
130 aa; identical to corresponding CDS from Y.pestis KIM5
PCD1 (EMBL:AF053946, AF074612) (130 aa), fasta scores;
opt: 861 z-score: 1608.5 E(): 0, 100.0% identity in 130 aa
overlap and to YERA_YERPE (EMBL:M34279) from Y.pestis
plasmid pYV019. Highly similar to TR:Q56910 (EMBL:Z18539),
syce, from Yersinia enterocolitica (130 aa) (99.28
identity in 130 aa overlap) and YERA_YEREN (EMBL:M4278),
yerh, also from Y.enterocolitica (130 aa) (98.5% identity
in 130 aa overlap). Similar to TR:Q51448 (EMBL:Z27629),
ORF1, Pseudomonas aeruginosa exoenzyme S ORF1 (116 aa)
(44.3% identity in 115 aa overlap)"
/codon_start=1
/transl_table=11
/label="syce"
/product="putative yope chaperone"
/protein_id="CAB54882.1"
/db_xref="GI:5832428"
/translation="MYSREDAITDLPOLSLSDPTIPYGVKVGACHTHPVQ
OIMFTPLSDNNDEKELLSHNIFSDILPILSWDEGVHPLMNKPLNSLDNNS
LYTQLEMLVQGAERLQTSLSLSPRSFS"
3600..4259
/gene="YPCD1.06"
3600..4259
/gene="YPCD1.06"
/feature="YPCD1.06"
/feature="YPCD1.06, yope, outer membrane vitulence protein,
len: 219 aa; identical to corresponding CDS from Y.pestis
KIM5 PCD1 (EMBL:AF053946, AF074612) (219 aa), fasta
scores; opt: 1409 z-score: 2140.1 E(): 0, 100.0% identity
in 219 aa overlap and to YOPA_YERPE (EMBL:M34279), YOPA,
from Y.pestis plasmid pYV019. Highly similar to YOPA_YERPS
(EMBL:U00543), yope, from Yersinia pseudotuberculosis
plasmid pIB1 (219 aa) (99.3% identity in 219 aa overlap)
and to YOPA_YEREN (EMBL:M92066), YOPA, from Yersinia
enterocolitica plasmid pYV (219 aa) (95.9% identity in 219
alignment_scores:
Quality: 1646.00 Length: 326
Ratio: 5.049 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.693
alignment_block:
US-08-699-716a-2 x YPCD1/rev ..
Align seg 1/1 to reverse of: YPCD1 from: 1 to: 70305
196 MetIleArgAlaTyrGlulnAsnProGlnHisPheIleGluAspLeuGI 212
|||||
22915 ATGATTAGAGCCCTACGACAAACCAACATTTATTGAGGATCTAGA 22866

```

```

212 ulysValArgValGlulnLeuThrGlyHisGlySerValLeuGIug 229
|||||
22865 AAAAGTTAGGTGGAAACAACTTACTGGTACGTCTTCTAGTTTAGAG 22816
|||||
229 lueValAlnLeuValLysAspLysAsnIleAspIleSerIleLysTyr 245
|||||
22815 AATTGGTCACTAGTCAAGAAATATAATATATATTCATTAATAT 22766
|||||
246 AspProArgLysAspSerGluValPheIleAsnArgValIleThrAspAs 262
|||||
22765 GATCCCAAGAAAGATTCGGAGCTTTTCCCATAGCATTAATACGATCA 22716
|||||
262 pIleGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThrI 279
|||||
22715 TATCGAATGCTCAAGAAATCCTAGCTATTTTCTACCCGAGATGCCA 22666
|||||
279 lLeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
|||||
22665 TCTTTAAAGCGCGTCATTATGACAACTCAAAATGGCATCAAGCA 22616
|||||
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTyrGluLeuArgAl 312
|||||
22615 GTRAAAGACTTCTCTCAATCATCGCGAATACCAATGGAAATTCGGGCG 22566
|||||
312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
|||||
22565 GTTCATGCGCAGTAATGCATTTCTTTAACCCCGCATCGTATGATGATG 22516
|||||
329 sPleLeuLysValIleValAspSerMetAsnHisHisGlyAspAlaArg 345
|||||
22515 AATATTTGAAGTGAATGATGTTGATCAATGAATCAATCATGATGATCCGCT 22466
|||||
346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyr 362
|||||
22465 AGCAAGTTCCGCGAAGAAATTAAGTCACTGATCCCGCAATTAAGATTTA 22416
|||||
362 rSerValIleGlnAlaGluIleAsnLysHisLeuSerSerGlyThrI 379
|||||
22415 TTCAGTTATTCACGCGAATTAATTAAGCATCTGCTAGTATGACACCA 22366
|||||
379 lAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
|||||
22365 TAAATATTCATATTAATCATTAATTCATGATGAATTAATTAATGAT 22316
|||||
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuGI 412
|||||
22315 TATACAGATGAAGACATTTTAAAGCCAGCCAGATACAAATCTCGA 22266
|||||
412 ulysMetProGlnThrThrIleGlnValAspGlySerGluLysIleVal 429
|||||
22265 GAAATGCGCTCAACCCATTCAGGTGATGGAGCGGAAATAATATG 22216
|||||
429 aLserIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
|||||
22215 TCTCGATTAAGACCTTCTTGAAGGAGAAATAAAGAACCGGGCGTGG 22166
|||||
446 GlyAsnLeuLysAsnSerLysSerLysAsnLysAspAsnAsnGluLeuSe 462
|||||
22165 GGTAAATCTAAAAACATCACTCTTATATTAAGTAAATTAATGAATTTATC 22116
|||||
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
|||||
22115 TCACCTTGCCACACCTGCTCGGATAAGTCCAGCGCGCTCAACGACTGG 22066
|||||
479 aLserGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
|||||
22065 TTAGCCAAAAAACAACTCACTGCTGATATTAATCAATCACTTTTATTC 22016
|||||
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetGI 512
|||||
22015 GCTATTAGAGCACTGACACCGCTTTCATTCAGAAATATGATTCAGTATGA 21966

```

512 nargleuensasparthserglylys 521
 |||||
 21965 ACGTCTGCTAGATGACACGCTGTGTTAAA 21938
 seq_name: gb_ba2:AF053946
 seq documentation block:
 LOCUS AF053946 70504 bp DNA circular BCT 06-OCT-1998
 DEFINITION Yersinia pestis plasmid pCD1, complete plasmid sequence.
 ACCESSION AF053946
 VERSION AF053946.1 GI:2996222
 KEYWORDS
 SOURCE Yersinia pestis.
 ORGANISM Yersinia pestis
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
 REFERENCE 1 (bases 1 to 70504)
 AUTHORS Hu, P., Elliott, J., McCreedy, P., Skowronski, E., Gaines, J., Kobayashi, A., Brubaker, R. R. and Garcia, E.
 TITLE Structural organization of virulence-associated plasmids of Yersinia pestis
 JOURNAL J. Bacteriol. 180 (19), 5192-5202 (1998)
 MEDLINE 98422474
 REFERENCE 2 (bases 1 to 70504)
 AUTHORS Hu, P., Elliott, J., McCreedy, P., Skowronski, E., Gaines, J., Kobayashi, A., Carrano, A. V., Brubaker, R. and Garcia, E.
 TITLE Direct Submission
 JOURNAL Submitted (16-MAR-1998) Biology and Biotechnology, Lawrence Livermore National Lab, L452, 7000 East Ave., Livermore, CA 94550, USA
 FEATURES
 source Location/Qualifiers
 1. /70504
 /organism="Yersinia pestis"
 /plasmid="pCD1"
 /strain="K1M"
 /db_xref="taxon:632"
 87. /1109
 /codon_start=1
 /transl_table=11
 /product="transposase"
 /protein_id="AAC62546.1"
 /db_xref="GI:2996223"
 /translation="MVFETVMEIKILHKGSSRAIAELGISRNTVRYLQAKSEP KYTPPPAVASLDEEDYRIQRIADAHYKIPATYIAREIDOGRGMTILRAEIR SLVPEQEPAYEFEREPRGROVOMGTMBNSPLHFAVAVIGSRMYITFTNMR YDTLECHRNAREFGVREVLYDMKATVYIQORDVYQGHREHSLMQPKMGFS PRLCRPRKQTKGKVRMYQVITNSFTILMLRLRWGITVDYETANKRGLWMLHDA NORHETIQARPCDRLWEQSKMLALPPEKKEYDALDENLVNFKHLHPHLSIYDS FCRGVA"
 1109. /1888
 /codon_start=1
 /transl_table=11
 /product="transposase"
 /protein_id="AAC62557.1"
 /db_xref="GI:2996234"
 /translation="MMELOHQRMLALAGOLESLISAAPALSOQAVDOEMDFLE HILHEKRLARHOKRQMTBMAFPATKFEEDPFPAGAPKOKOSRSLSTERN ENIVLPGSGVGTHTLAIMGIEYAVAGIKVPTTADLLDLSIQRGGRKTTIQR GWAPRLILDEIGYLPFSQEEKLFFQVIARYESAMILTSNLPFGQMDPTFAGDA ALTSAMLDRIHLHSHVQIKGESYRLRKRKAGVIAENPE"
 complement(1939. /2343)
 /codon_start=1
 /transl_table=11
 /product="transposase"
 /protein_id="AAC62568.1"
 /db_xref="GI:2996245"
 /translation="MTEYQASERRCGRINGISRLHYCPNTARDIPVVELOKLAHQ YPAGFGLMNLROSGLPNNVRYRYRLKLINRRGRKRLPRHQPPLAIPKMK NHCWYSDMSDLTDGRRRRLVVEILFVINDG"
 complement(2379. /2645)
 /gene="lcrs"
 complement(2379. /2645)
 /gene="lcrs"
 CDS
 gene
 CDS

/codon_start=1
 /transl_table=11
 /product="low calcium response locus protein s homolog"
 /protein_id="AAC62579.1"
 /db_xref="GI:2996256"
 /translation="MKKARFTETQILRYLAKVEGGRHKYDRENGSEASYMKKSK TGGMSDICKMKKEREENRRKQMTASLSIDHETLADVAKKL"
 complement(3193. /3540)
 /gene="yscM"
 complement(3193. /3540)
 /gene="yscK"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein M homolog"
 /protein_id="AAC62586.1"
 /db_xref="GI:2996263"
 /translation="MKINTLOSILNQOITQVGHGGQAGRLTETNPENSHQISTAEK AFANVLEHVNTALSRHDICLPRVSNELKQKAGGVITVGTREQLSDPAKLL LEAARQOTADG"
 complement(3765. /4430)
 /gene="yscL"
 complement(3765. /4430)
 /gene="yscI"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein L homolog"
 /protein_id="AAC62596.1"
 /db_xref="GI:2996273"
 /translation="MSQTCQGVYAMPFVQVITPSNLSLACGLILRAEDYQSITTE ELISAQDASKILADAEVEYEQQKLOMGOMGEARTLQTLIHTOLOQOQFYRHV EQOMSEVVLAVRKILNDQVAMTLOVRAVLVNSKQGVAVRVNDAVGALREYI AKVHKDFEISYLEVTADARLDQGCILETVGIIDASIDQIALSRNISTITGQMK VTE"
 complement(4376. /5005)
 /gene="yscK"
 complement(4376. /5005)
 /gene="yscK"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein K homolog"
 /protein_id="AAC62605.1"
 /db_xref="GI:2996282"
 /translation="MKNYITSPFLRCPAAVYILHLEQPSLMSRSLPYLPQWRDASHL NNALEDTIDVEEPHGLGALPQPOSQELICRLGVLHGEAIRCVLASPIQO LLTYNQSTLQIIVQHGLLIGPNPTHQRLPRIEBRTYIQSGIAFWLAAMEPOQ AMCKNLSTRLPLATSEFWLVAESQRLAQTLCHLVQVPTCSHLEK"
 complement(5005. /5739)
 /gene="yscJ"
 complement(5005. /5739)
 /gene="yscJ"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation I1protein J"
 /protein_id="AAC62607.1"
 /db_xref="GI:2996284"
 /translation="MKVTSSTLILILFLTGCYVDLYTGISQKGNEMALALQOGL SADKPPDQSGIKILVSESDVAQADIDIKRKGYSSTLQDVPKQGLSSPIEL ARLNARAKQESLRLSETDGLVAVRVLPEEDNNKRRKVAASVFIHADIQF DTYLPIQIVQVNNNSJEGLYADRIIVYLTVPSVYROSSHLPRNTSILTSIOVSESGHL IGLSILLPLPVTLAQYFWLQRRK"
 complement(5746. /6093)
 /gene="yscI"
 complement(5746. /6093)
 /gene="yscI"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein I homolog"
 /protein_id="AAC62608.1"
 /db_xref="GI:2996285"
 /translation="WPNTEIAQDEVIITLTLEGPAPETTDQIMRDAANSEDTQGL GHSLEKVEDIOKSEFTYKSDLHTRKLAIVSDNPDLMLQMSLIRITIOEBLIAKTAG RMSQNVETLSKGG"
 complement(6094. /6591)
 gene

/gene="ysch"
 complement(6094. .6591)
 /gene="ysch"
 /codon_start=-1
 /transl_table=-1
 /product="Yop proteins translocation protein H homolog"
 /protein_id="AAC62547.1"
 /db_xref="GI:2996224"
 /translation="MTVTLKRGSTITLSSSSQAVSTLQPYASLKTQLENKKSESAB
 KTRFVLMQDYASNPDPDAVLEATPVPALARGQHGQSVPAIDLPELRSVLAQ
 FDSFGKRWBAIILQVLEGIKPNESQVGLPYSELINKELMILPNSIYDSLHNSHO
 IDMDT"
 complement(6588. .6935)
 /gene="yscg"
 complement(6588. .6935)
 /gene="yscg"
 /codon_start=-1
 /transl_table=-1
 /product="Yop proteins translocation protein G homolog"
 /protein_id="AAC62548.1"
 /db_xref="GI:2996225"
 /translation="MKTKLVLAETALIGTGNHYHEANCIAMWHLKGEERAVOLI
 RLSSLNKRGDYASALQOGNKLAAYPDLEPMIALCEYRLGIGSALESRLNRLARSQDPRI
 QRFVNGMRBQLKT"
 complement(6937. .7200)
 /gene="yscf"
 complement(6937. .7200)
 /gene="yscf"
 /codon_start=-1
 /transl_table=-1
 /product="Yop proteins translocation protein F homolog"
 /protein_id="AAC62549.1"
 /db_xref="GI:2996226"
 /translation="MSNFGFTKGTDIADLDAVAQTLKRPADANKAVNDSIALKDK
 PNPALADLQHSINRMSVYININSTIVSMKMDLQGILOKEP"
 complement(7201. .7401)
 /gene="ysce"
 complement(7201. .7401)
 /gene="ysce"
 /codon_start=-1
 /transl_table=-1
 /product="Yop proteins translocation protein E homolog"
 /protein_id="AAC62550.1"
 /db_xref="GI:2996227"
 /translation="KTQLEQHLNVTWSTINQLEMAITLTKKDMRGADAKQYQVW
 QRESKALSAIAIIRYVAGDLK"
 complement(7398. .8657)
 /gene="yscd"
 complement(7398. .8657)
 /gene="yscd"
 /codon_start=-1
 /transl_table=-1
 /product="Yop proteins translocation protein D homolog"
 /protein_id="AAC62551.1"
 /db_xref="GI:2996228"
 /translation="MSWVCRFYQGRHGVVEVLPKRCYFGSDPLQSDIVLSSEIAP
 VHYLVWDEGRITDLSAEPLDEGLPVPGLTLKAGSCLEVGFLMTVAAGQPIPE
 TLQVPTQRKEPTDLRPSRLGIGVLSLLLTFLGLHGHMREYNDDQGVDEQEV
 RLRLATAAKDVVLTSPKKEGEPMLTGYIODNHARLSIONFLESHGIFPRLERAME
 ELRQGAFFLQRLQRIEYSLAPQAGMTOLNVESEEBEOKRISDLAEVYGLGIV
 ESKVRIAGNQRRKLDALEQGLDSDFTYVVKELIELNGOVNDEKLNFPNLOQCFR

alignment_scores:

Quality: 1646.00 Length: 326
 Ratio: 5.049 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.693

alignment block:

US-08-699-716a-2 x AF053946 ..

Align seg 1/1 to: AF053946 from: 1 to: 70504

196 Metleatgataatyruglunasp3proclnhsphelleglunasp1eugl 212

25311 ATGATTAGCCTTACGACCAACACCATTTTATGTGGAGTCTAGA 25360
 212 ULyVaLaRgYaIguIneUthrGlyHsGlySerSerValleuG 229
 25361 AAAAGTTAGGTGGACACTTACTGGTCAGTCTTCTAGTTTAAAG 25410
 229 lUeUvAlGhLeUvAllySaPlYsaSnIlleAsPlleSerlleYsTyr 245
 25411 AATTGCTCACTTACGATAAAGATAATATATATTCATTAATAATAT 25460
 246 AsPrOaRdLySaSpSerGluVAlPhaAlaSaRgYalIleTrAsPaS 262
 25461 GATCCAGAAAGATTCGAGAGTTTTCACATAGAGTAATTAAGTGA 25510
 262 pIleGluLeUleUlySylleUleAlaTyrPhaLeUProGluAsPthri 279
 25511 TATGAATGCTCAAGAAATCCTAGCTTATTTCTACCGAGAGATGCCA 25560
 279 lElEuLyGlyGlyHsTyrAsPaSnGluIneUgInaSnGlylleYsaRg 295
 25561 TTCTTAAAGCGGCTCATATATGACCAACACGCAAAATGGCATCAACGA 25610
 296 ValLySGluPhaLeUgInuSerSerProAsnThriUleUaRgAl 312
 25611 GTAAAGAGTTCCTTGATCATCGCCGAATACACAAATGGGAATGGCGG 25660
 312 aPhMeLaVaImeThsPheSerleUthrAlaAsPaRgIlleAsPaS 329
 25661 GTTCATGGCATAATGCAATTCCTTTAACCGCGGATCGATCATATG 25710
 329 sPlleUlySvAlIleYalAspSerleAsnHsHsGlyAsPaLaRg 345
 25711 ATATTTGAAGATGATTGTCATCAATGATCATCATGATGATGCCGT 25760
 346 SerLyLeuAArgGluGluLeuAlaGluLeuThraGluLeuYsIleY 362
 25761 AGCAAGTTCCTGAAGATTATGACCTTACCCGCAATTAAGATTGA 25810
 362 rSerValIleGlnAlaGluIleAsnLyHsLeuSerSerGlyThri 379
 25811 TTCAATGATTAACCCGAATTAATATAGCATCTCTAGTAGTGCCACA 25860
 379 lEaSnIlleHsAsPlYsSerlleAsnLeUeTasPlYsaSnleUtyrGly 395
 25861 TAAATATCCATGATTAATCCATTAATCTCATGATTAATAATTAATG 25910
 396 TyrThraSpGluGluIlePhaLyAlaSerAlaGluUtyrYsIleUgl 412
 25911 TATACAGTAGAAGATTTTAAAGCCAGCCAGATCAAAATTCGCA 25960
 412 ULyMeTProGlnThrThrIleGlnValaSpGlySerGluYsYsIleY 429
 25961 GAATATGCTTAACCAACCATTCAGTGGATGGAGGAGAAAAAATAG 26010
 429 aISerlleYsaSpHleUnglySerGluSnuYsArGthrGlyAlaLeu 445
 26011 TCTGATAAAGAGACTTCTTGAAGTAGAATAAAGAACCGGGCGCTTG 26060
 446 GlySnuLeuYsaSnSerYrSerYrAsnLySaSpSaSnGluLeuSe 462
 26061 GGTATCTGAAGAACTCATCTTATTAATTAAGATATATATGATATATC 26110
 462 rHsPhaAlaThrThrCySsaSpYsSerArGProLeuAsnAspLeuY 479
 26111 TCATTTGCCACACCTGCTCGATAGTCCAGGCCCTTAACGACTTGG 26160
 479 aISerGlnLyThrThrGlnLeuSerAspIleThrSerArGpHeaSnSer 495
 26161 TTACCAAAAAACAATCAGCTGCTGATATATCATCATGCTTTAATCA 26210
 496 AlAtIleGluAlaLeuAaRgPheIleGlnYsTyrAspSerValMeTgl 512

26211 GCTATTGAAGCCTGACCGTTTCATTCAGAAATGATCAGTATGCA 26260

512 nargleuenaaspthrserglylys 521
 26261 ACCTCTGCAGATGACGCTGTGTA 26288

seq_name: gb_ba2:AF074612

seq_documentation_block:

LOCUS AF074612 70559 bp DNA circular BCF 07-Apr-2000
 DEFINITION Yersinia pestis plasmid pCD1, complete plasmid sequence.

ACCESSION AF074612 M25810

VERSION AF074612.1 GI:3822037

KEYWORDS Yersinia pestis.

SOURCE ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Yersinia.

1 (bases 4318 to 44840)
 Leung, K.Y. and Straley, S.C.

The yopM gene of Yersinia pestis encodes a released protein having

homology with the human platelet surface protein GPIIb alpha

J. Bacteriol. 171 (9), 4623-4632 (1989)

2 (bases 1 to 70559)
 Perry, R.D., Straley, S.C., Fetherston, J.D., Rose, D.J., Gregor, J. and

Blattner, F.R.

DNA sequencing and analysis of the low-Ca²⁺-response plasmid pCD1

of Yersinia pestis KIM5

Infect. Immun. 66 (10), 4611-4623 (1998)

3 (bases 4318 to 44840)
 Straley, S.C.

Direct Submission

Submitted (26-APR-1993) Microbiology and Immunology, University of

Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA

4 (bases 1 to 70559)
 Perry, R.D., Straley, S.C., Fetherston, J.D., Rose, D.J., Gregor, J. and

Blattner, F.R.

Direct Submission

Submitted (25-JUN-1998) Microbiology and Immunology, University of

Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA

On Apr 7, 2000 this sequence version replaced gi:155539.

Location/Qualifiers

1..70559

/organism="Yersinia pestis"

/plasmid="pCD1"

/strain="KIM5"

/db_xref="taxon:632"

57..368

/gene="Y0001"

/note="O103; 43 pct identical (0 gaps) to 100 residues of

an approx. 200 aa protein GENPEPT: gi1537126,orf_O198

Escherichia coli"

/codon_start=1

/transl_table=11

/product="unknown"

/protein_id="AAC69758.1"

/db_xref="GI:3822038"

/translation="MHQSRGAAGSRTSLRMROSGYVYVWMLARRLMRCGLASROP

GKPRYRGREVSASPDLKRFKRPSEPNVWSGIYIKVNGWCYLAIVIDLYSRH

W"

665..1033

/gene="nuc"

/note="Y0002"

665..1033

/gene="nuc"

/codon_start=1

/transl_table=11

/product="endonuclease"

/protein_id="AAC69759.1"

gene

CDS

1171..1425

/gene="repB"

/note="Y0003"

1171..1425

/gene="repB"

/codon_start=1

/transl_table=11

/product="replication protein"

/protein_id="AAC69760.1"

/db_xref="GI:3822040"

/translation="MSQIENAVTSSSKRAYRGKGNPLTGAEKQMSYSRKKEHKAINV

FIQNDIKNFIQCEGTCGTQETEMIRWQREKARTNA"

complement(1560..1643)

/note="antisense RNA"

/product="copA"

1667..1741

/gene="tap"

/note="Y0004"

1667..1741

/gene="tap"

/codon_start=1

/transl_table=11

/product="repa translation protein"

/protein_id="AAC69761.1"

/db_xref="GI:3822041"

/translation="MRKVQYLLRLLLPCNISGRCD"

1734..2600

/gene="repa"

/note="Y0005"

1734..2600

/gene="repa"

/codon_start=1

/transl_table=11

/product="replication protein"

/protein_id="AAC69762.1"

/db_xref="GI:3822042"

/translation="MTNHQALFTTHIRYQKNPNEPTREGEKTLPECKRLAKAEGF

TSRFDLSMVAFAFSLSRHMPPLRRRAIDALOGMCFHDPPLANRISITNIAI

ECGLATESGVSAGNISITRATRALFELSGLIYQVYDQICNIPITDITFPALFA

LDVSDVAVAARSRVEMQOEROKRLPRLDEDLIAAMRFRFRPSYQTERAAH

GLRARARDDVDRTRDIAIVNQRLRIAGRFNGINDAVRERARVKEKMLSR

NNYTRLATGAT"

1741..2935

/note="oriR"

complement(3427..3645)

/gene="Y0006"

complement(3427..3645)

/gene="Y0006"

/note="O15 (f72); 42 pct identical (0 gaps) to 33

residues of an approx. 216 aa protein GENPEPT: gi12055297,

Imp2 Xenopus laevis"

/codon_start=1

/transl_table=11

/product="unknown"

/protein_id="AAC69763.1"

/db_xref="GI:3822043"

/translation="MRSPLAYGGSVYTVVYQQLHNDIKSAVKHKRLVHVLGQASS

AOVFSPHPRGMAAPADIVNSVDNRR"

4758..5186

/gene="Y0008"

4758..5186

/note="O17 (o142); 31 pct identical (1 gap) to 48

residues of an approx. 104 aa protein GENPEPT: gi12149940,

orf1 p. syringae"

/codon_start=1

/transl_table=11

/product="unknown"

/protein_id="AAC69764.1"

/db_xref="GI:3822039"

/translation="MDTKLQHTPIGTWVDRPVNTKSGKRLRCPDFVHYRMDL

VNAGIPVFTVNSFKALHDKVITVDGKNTQMGSEFNSQAQVQNSSEVLLIMDDFTVQ

AYLVQWSPRNKNGKDWRSY"

1171..1425

/gene="repB"

/note="Y0003"

1171..1425

/gene="repB"

/codon_start=1

/transl_table=11

/product="replication protein"

/protein_id="AAC69760.1"

/db_xref="GI:3822040"

/translation="MSQIENAVTSSSKRAYRGKGNPLTGAEKQMSYSRKKEHKAINV

FIQNDIKNFIQCEGTCGTQETEMIRWQREKARTNA"

complement(1560..1643)

/note="antisense RNA"

/product="copA"

1667..1741

/gene="tap"

/note="Y0004"

1667..1741

/gene="tap"

/codon_start=1

/transl_table=11

/product="repa translation protein"

/protein_id="AAC69761.1"

/db_xref="GI:3822041"

/translation="MRKVQYLLRLLLPCNISGRCD"

1734..2600

/gene="repa"

/note="Y0005"

1734..2600

/gene="repa"

/codon_start=1

/transl_table=11

/product="replication protein"

/protein_id="AAC69762.1"

gene /db xref="GI:3822044"
/translation="MINTPTFELLPKIAHREGLKLSODEYGLCELLINDRYVIMLAR
DEILNRLLTLGLFSCGPEARSASQULFFCYISIALNKDGPCTRAWSEELGLINFKH
SIDLENVENSKEIANFYDWLSVLSLPAETAPSYISIG"
5204. 7402
/gene="ypka"
/note="Y0009"
5204. 7402
/gene="ypka"
/codon_start=1
/transl_table=11
/product="secreted protein kinase"
/protein_id="AAC69765.1"
/db_xref="GI:3822045"
/translation="KSYKIMGTNPPSISLAKAHERISQHQNVYGEINIGKRRITII
DNQVFLNPHSGESLFRSGVGRIFSGKFNFSIANLTDTLAAOKTTSQLSRSDIN
ALSNFLEKAPQTEPLGKMGGEPLSGAPDLEMRVAETDKFAGEGSHSIITKQKORR
VAKIRRSIAEGHLFAELRAVKIHYTACKHFNLAHVGMNAVYVAGNREKDALNDEYVA
GRCSDITLADSKMGKOGINSEAVWGTFIKFLAHLIDVTNHLKAGVANDIDRGWNV
VDRASGEPVIVDGLIHSRGCPKPFESKAPICVNGISAGEASQVPLVYSTLHL
CIEGEBKNEPIPNQGLRITSEPAHVDENGTPIHRGILNGVETAKETRIITDILGVS
ADSRDPSNARLHEFLSDGTIDEEBARKILKDTVLGENSPJSTOVRRTTRPKLELSD
ILRTHSGSATKQOLDMGVLSDDLTMVALDKAEREHGGVDDKQSFNSILKTVRYVY
EDYVGRREDTQSESTEVSFYRNSFMLSIYEPESLQRIOKHLDQTHSPSDLSVRAH
KLELLELVTLISQOOGPVSVEYTGFLNRLTEATITISQOLNTLQOOGESAKOLSI
LINRSGSNADVAROSLQRRDSQPYVKEGTQDYTLIHQMAAHARITLQGVSEFTDD
MNFVYDSLEPLIQLGRSLMDHELVDEORERKLETTIAERLNLEREMW"
7798. 8664
/gene="yopJ"
/note="Y0010"
7798. 8664
/gene="yopJ"
/codon_start=1
/transl_table=11
/product="targeted effector protein"
/protein_id="AAC69766.1"
/db_xref="GI:3822046"
/translation="MIGPISQINISGGISEKETSLSINEELKNITQLEPDISDGS
FKRNSRDVEMPALVTDANKKIPENMLNLTSLDLSIEIKVNIENGVSRRITIN
MEGGHIEVDYKIRHNGTSLILPEPNNFNSMGAMLAIFKTAIEYQLPDPHFS
VMDQIORSSECCIGIFSLAKLKYLERPSLLKIHNDNIGLISGENTLPDKDLPYV
PYTFKHQIGKRNLEYINTNPGVGVVANKNETIYVNRFNKNSIYVGKELSYVHR
KIAEKETILKY"
9858. 10064
/gene="Y0012"
9858. 10064
/gene="Y0012"
/note="068: 45 pct identical (0 gaps) to 68 residues of an
approx. 560 aa protein GEMPBT: g112182435, hypotheticala1"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC69767.1"
/db_xref="GI:3822047"
/translation="NAVGRKNMLFAGSLRAGQRMASILSLLETKAKNGHDYVYLROV
LTRLPLWNSQNALPFAENRFS"
10347. 11753
/gene="yopH"
/note="Y0013"
10347. 11753
/gene="yopH"
/codon_start=1
/transl_table=11

```
alignment_scores:
  Quality: 1646.00      Length: 326
  Ratio: 5.049          Gaps: 0
Percent Similarity: 100.000  Percent Identity: 99.693
```

```
alignment_block:
US-08-699-716A-2 x AF074612 ..
```

Align seg 1/1 to: AF074612 from: 1 to: 70559

196	MetIleArgAlaTyrGluUcIlnasnProGlnHisPheIleGluAspLeuG1	212
37919	ATGATTAGAGCCTACGACAAACCCACACACTTTTATTAGAGACTCTACA	37968
212	ULYSVALARVALGIUGInLeuThrGlyHisGlySerSerValLeuGluG	229
37969	AAAAGTTAGGGTGAACAACCTACTACGGTCATCGTCTTCAGTTTAAAG	38018
229	LULeuValGlnLeuValLysAspLysAsnIleAspIleSerIleLysTyr	245
38019	AATTGGTTCAGTTACTCAAAAGTAAATAATATGAAATATTCATTAAATAT	38068
246	ASPProArgLysAspSerGluValPheIleAsnArgValIleThrAspAs	262
38069	GATCCCAAGAAAGATTCGGAGGTTTTTGCCAAATAGATTAATCTGATGA	38118
262	PIleGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThrI	279
38119	TATCGAATTGCTCAAGAAATCCTAGCTTATTTCTACCCGAGATGCCA	38168
279	LELeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg	295
38169	TTCCTTAAAGCCGTCATTATGCAACCAACTCAAAATGGCATCAAGCA	38218
296	VALYsGluPheLeuGlnSerSerProAsnThrGlnTPGInLeuAspAl	312
38219	GTAAAGAGATTCCTTGAACTACATCGCGAATATACATATGGAAATTCGG	38268
312	APhMetAlaValMetHisPheSerLeuThrLaaAspArgIleAspAspA	329
38269	GTTCATGGCAGTAATAGCATTTCTTTAAACCCGATGCTATCATGATAG	38318
329	SPILeuleuLysValIleValAspSerMetAsnHisHisGlyAspAlaArg	345
38319	AATATTTGAAAGTGAATGTTGTGATTCAATCAATCAGATCGTATCCCGT	38368
346	SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyr	362
38369	ACCAAGTTCGGAGAAATTTACTCGCTTACCGCGAATTAAAGATTTTA	38418
362	rSerValIleGlnAlaGluIleAspLysHisLeuSerSerGlyThrI	379
38419	TTCAGTATTCAACCGGAATTAATATAGCATGTGCTATATATGGCACCA	38468
379	IleAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly	395
38469	TAAATTCATCATATAAATCCATTATCTCATGATTAATAATTTATATGCT	38518
396	TyRThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuG1	412
38519	TATACGATGAAGAGATTTTAAACCCACGCGAGATGAATAAATCTCGA	38568
412	ULysMetProGlnThrThrIleGluValAspLysGlyLysIleVal	429
38569	GAAATAGCCTCAACACCATTCAGGTGATGGAGCGGAAAAAATATG	38618
429	ALSerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu	445
38619	TCTCGATTAAGACCTTCTTGGAAGTGAAGATAAAGAAACCGGGCGTTG	38668
446	GLysAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe	462
38669	GGTAATCTGAAAAACATCACTACTTATATATAAAGATTAATAAGATTATTC	38718
462	rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV	479
38719	TCACTTGGCACCACTGCTCGGATTAAGTCCAGGCGCTCAACGACTTGG	38768
479	ALSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer	495
38769	TTAGCCAAAAAACACTACGCTGCTGTGATATTACATCAAGCTTTTATATCA	38818

456 AlaAIEGUALIleuAsnArgPheIleGlnLYsTYrAspSerValMetG1 512
38819 GCTATTGAAGCCGTGAACCGTTTCATTTCGAAATATGATTCATGTATGCA 38865
512 nArgLeuAspAspPThrSerGlyLys 521
38869 ACGTGTGCTAGCTGACACGTCCTGGTAA 38896

seq_name: gb_pat:A46411

seq_documentation_block:

LOCUS	1014 bp	DNA	PAT	07-MAR-1993
A46411				
DEFINITION	Sequence 1 from Patent WO9524475.			
REMARKS				

SOURCE	Yersinia pestis
ORGANISM	Yersinia pestis

Yersinia.

COMMENT	Other publication AU 1853395 950925
FEATURES	Location/Qualifiers
SOURCE	1..1014

CDS

BASE COUNT	346 a	181 c	201 g	286 t
ORIGIN				

```
alignment_block:
    no 00 000 715
```

US-08-699-716A-2 x A46411

Align seg 1/1 to: A46411 from: 1 to: 1014

194 GlPheMetLeIerGlaIaTyrgluGluInaAnProGlnHISpHeIleGluAs 210
7 GAATTC...ATTAGACCTTACGACACAAAACCCACAACATTTATTAGAGA 53
210 pleuGluLysValaIarGvaIGluGluIneuThrGlyHISGlySerValL 227
54 TCTGAAAAAGTTAGGCGTGGACCAACTTACTGTCATGTTCTTCAGTTT 103
227 eugGluGluIneuValGluIneuValLysAspLysAsnLLeaspLLeSerLe 244
104 TAGAAGATTGGTTCAGTTAGTCAAGAGATAAAAATATGAAATTTCCATT 155
244 LysTyIAspProArGlyAspSerGluValPheIaAsnArGValIleTh 260
154 AAATATGATGCCGAAAGATTTCCGAGGTTTTTGGCATATAGGTATTTAC 207
260 rAspAspLLeGluIneuLysLysLLeuIaLaTyIrrPheIeuProGluA 277

Kantzen

```

204 TGATGATATCGAATTGGCTCAGAAATAACCTTAGCTATTTCCTACCCGAGG    253
277 GETHrAlEuleuylsglycylHisTyrAspAsnGlnLeuGlnAsnGlyIle    293
254 ATGCCAATCTTTAAAGGCGCGCATATTATGCACAACCNACTGCAAAAATGGCATC    303
256 lysArgVallylsGluPheLeuGlnSerSerProAsnThrGlnTrpGlu     310
304 AAGCAGAGAATAAGAGTTCCTTGGAATCATCGCCGGAATACAAATGGGAATT    353
310 uArgAlaPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleA     327
354 GCGGGCGGTTCATGCGCAGTAATGCATTTCTCTTTAACCCGCGATCGTATCG    403
327 spspAspDileuLysValIleValAspSerMetAsnHisHisGlyasp     343
404 ATGATGATATTGTAAGTAGATGTGTTGATTCGAATGAATCATCATCANGGAT    453
344 AlaArgserLylysLeuArgGluGlnLeuAlaGluLeuThrAlaGluLeuLy   366
454 GCCCTTACCAAGTTGCGGGAAGAATTAGCTGACTGATCCCCCGAATTAAAA    503
350 sileYrservallieGlnAlaGluIleasnlyHisLeuSerSerg       377
504 GATTATTACAGTATTATCAAGCCGAATAATTAATACCATCTGCTAGTAGTG    553
377 lythrIleasnIleHisAspLysserIleasneumetAspLyAsnLeu     393
554 GCACCATAAATATCATGATATAATCCATTAATCATGATATAAAATTTA      603
394 TyrglyTyrrhrAspGluGlnIlePheLysAlaSerAlaGutryLysIl     410
604 TATGGTATTACAGATGAAGAGATTTTAAACCCAGCCAGCATGTCANAAT     653
410 eleuGlnLysmetProGlnThrThrIleGlnValAspGlyserGlusyl     427
654 TCTCGAGAAATGCGCTCAAMACCATTACAGTGGAGGGGCGCAGAAAA      703
427 ysIleValSerIleLysAspPheLeuGlyserGluAsnLysArgThrGly     443
704 AAATAGTCTCGATTAAGACATTTCTTGAGACTGAGAAATTAAGAACCGGG    755
444 AlaleuGlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnaEngl    460
754 GCGTTGGGTAAATCGAAAACTCATACCTTATATAAAGATATATATGA      803
460 uLeuSerHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnA     477
804 ATTATCTCACTTGGCACACCTCGCTCGGAATAATCCAGGCGCGCTCAAG    853
477 splenValSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPhe     493
854 ACTTGTTAGGCCAAAAACAACACTAGCTGTGTGATTATACATCCGTTTT     903
494 AsnSerAlaIleGlnAlaLeuAsnArgPheIleGlnLysTyrAspSerVa     510
904 AATTAGCTATTGAAGCCTAAGACGTTTCATTAGAAATATGATTCACT      955
510 lmetGlnArgLeuAsnAspAspThrSergLylys      521
954 GATGCACGCTGCTGATGACAGCTGCGTAA      967
seq_name: gb_pat:A56793

```

seq_name: gjb_pat:A56793

seq_documentation_block:

LOCUS	1014 bp	DNA	PAT	03-MAR-1998
A56793				
DEFINITION	Sequence 1 from Patent WO9628551.			

SOURCE	Yersinia pestis
ORGANISM	Yersinia pestis

[illegible]

seq_name: gb_ptt:A66413

seq_documentation_block:

LOCUS	A66413	1014 bp	DNA	PAT	07-MAR-1997
DEFINITION	Sequence 3 from Patent W09524475.				
ACCESSION	A66413				
VERSION	A66413.1 GI:2300614				
KEYWORDS					
SOURCE	Yersinia pestis.				
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.				
REFERENCE	1 (bases 1 to 1014)				
AUTHORS	Tieball,R.W., Williamson,E.D. and Leary,S.E.				
TITLE	VACCINE COMPOSITIONS				
JOURNAL	Patent: WO 9524475-A 3 14-SEP-1995;				
COMMENT	SECR DEFENCE BRIT (GB)				
FEATURES	Other publication AU 1853995 950925.				
SOURCE	Location/Qualifiers				
	1..1014				
	/organism="Yersinia pestis"				
	/db_xref="taxon:632"				
	1..987				
	/note="named protein product; Protein sequence is in conflict with the conceptual translation"				
	/codon_start=1				

/transl_table=11
 /protein_id="CA02894.1"
 /db_xref="GI:2300615"
 /translation="GIRIRAYEONPOHFIEDEKVEVLEQTHGSSVLEIYOLVND
 KNIDISIKYPRDRSEYFANRVITDDIELKILATYLPEDAILKGGHYNOQNGIK
 RYKFELESSPNTOWELRAFAVYHFSITADRIDDIUKYVDSNNHGDARSKIREL
 AELTAEIKITSVIOAELINKHLSSTGTTINIDKSLNMDKNLYGTDEIFKASAEYKI
 LEKMPOTTIOVDSEKIKVISIKDFLSENNRTALGNLKNYSYNDKNDELSPHATTC
 SDRSRPLNDIVSOIKTQLOSDITSFRNSAIBALNRFIOKYSVMQRLDDTSRK"

BASE COUNT 343 a 184 c 205 g 282 t
 ORIGIN

alignment_scores:
 Quality: 1641.00 Length: 325
 Ratio: 5.049 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.692

alignment_block:
 US-08-699-716A-2 x A46413 ..

Align seg 1/1 to: A46413 from: 1 to: 1014

```

197 ILeatgalaTyrGlulGlnAsnProGlnHisPheIleGluAspLeuGlu 213
13 ATTCGAGCCTACGAAACCAACCAACATTTATGAGGATCTAGAAAA 62
213 sValaArgValGluGlnLeuThrGlyHisGlySerValLeuGluGlu 230
63 AGTTAGGGTGAACACACTTACTGCTATGCTTCTCAGTTTGAAGAAAT 112
230 euValGlnLeuValLysAspLysAsnIleAspIleSerIleLysTyrAsp 246
113 TGGTTCAGTACTCAAGATAAAAAATAGATATTTCCATTAATATGAT 162
247 ProAtgLyAspSerGluValPheAlaAsnArgValIleThrAspAsp 263
163 CCCAGAAAGATTCGAGGAGTTTTCGCAATAGAGTAATATGATGATAT 212
263 eGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGlnHisPheIle 280
213 CCAATTTGCTCAAGAAATCCTAGCTATTTTCCACCGAGAGAGCCATTC 262
280 euLysGlyLysIleTyrAspAsnGlnLeuGlnAsnGlyIleLysAspVal 296
263 TTAAGGCGGCTCATTTGACCAACCACTGCAAAATGGCATCAAGCAGTA 312
297 LysGluPheLeuGluSerSerProAsnThrGlnTrpGluLeuArgLap 313
313 AAAGAGTTCCTGAATCATCGCCGGAATACACAATGGGAATGGCGGCTT 362
313 eMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAspAsp 330
363 CATGGAGAAATGCAATTTCTTTAACCCCGATCCTATCGAAGAGARA 412
330 ILeuLysValIleValAspSerMetAsnHisGlyAspAlaAspSer 346
413 TTTTGAATGATGTTGATTCATCAATCATCATGATGATGCCCTAGC 462
347 LysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyr 363
463 AAGTTCGCGGAGAAATTAAGTTCAGCGGCAATTAAGATTTATATTC 512
363 rValIleGlnAlaGluIleAsnLysHisLeuSerSerSerGlyThrIle 380
513 AGTTATTCAGCGGAAATTAAGATCTGCTAGTAGTGCGACACATA 562
380 snIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGlyTyr 396
563 AATTCATCATTAATCATTAATCATGATTAATTAATGATGAT 612
397 ThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuGlu 413

```

```

613 ACAGATGAAGAGATTTTAAAGCCAGCCAGATACAAATTCGAGAA 662
413 sMetProGlnThrThrIleGlnValAspGlySerGluLysIleVal 430
663 AATGCTTAACCAACCATTCAGTGGATGGAGCGAGAAAAAATAGTCT 712
430 eIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeuGly 446
713 CATAAAGAGACTTCTTGAAGTGAATTAAGAAACCGGGCGGTGGGT 762
447 AsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSer 463
763 AATCTGAATAACTCATCTATATATTAAGATTAATATGATATCTCA 812
463 sPheLeuThrThrGlySerAspLysSerArgProLeuAsnAspLeuVal 480
813 CTTTCCACACCTCTCTCGATTAAGTCCAGCGGCTCAACGACTTGITA 862
480 eGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSerAla 496
863 GCCAAAAAACCACTCAGCTGTGATATTAATCACTGTTTATACACT 912
497 ILeuAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetGlnArg 513
913 ATTGAAGCACTGACCGTTTCATTCAGAAATATGATCAGTATCAACG 962
513 gLeuLeuAspAspThrSerGlyLys 521
963 TCTGCTAGATGACGCTGTGTA 987

```

seq_name: gb_pat:A56795

seq_documentation_block: 1014 bp DNA PAT 03-MAR-1998
 LOCUS A56795
 DEFINITION Sequence 3 from Patent WO9628551.
 ACCESSION A56795
 VERSION A56795.1 GI:3712810
 KEYWORDS
 SOURCE
 ORGANISM
 Yersinia pestis.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.

REFERENCE
 1 (bases 1 to 1014)
 TITLE Tiltball R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C.,
 AUTHORS Bennett and Alice, M.
 JOURNAL VACCINES FOR PLAGUE
 PATENT: WO 9628551-A 3 19-SEP-1996;
 COMMENT SECR DEFENCE (GB)
 OTHER PUBLICATION ZA 9602036 960716
 OTHER PUBLICATION AU 4951196 961002.

FEATURES
 source
 1..1014
 location/Qualifiers

CDS
 1..990
 /db_xref="taxon:632"
 /note="unnamed protein product"

/codon_start=1
 /transl_table=11
 /protein_id="CA03420.1"
 /db_xref="GI:3712811"
 /translation="GIRIRAYEONPOHFIEDEKVEVLEQTHGSSVLEIYOLVND
 KNIDISIKYPRDRSEYFANRVITDDIELKILATYLPEDAILKGGHYNOQNGIK
 RYKFELESSPNTOWELRAFAVYHFSITADRIDDIUKYVDSNNHGDARSKIREL
 AELTAEIKITSVIOAELINKHLSSTGTTINIDKSLNMDKNLYGTDEIFKASAEYKI
 LEKMPOTTIOVDSEKIKVISIKDFLSENNRTALGNLKNYSYNDKNDELSPHATTC
 SDRSRPLNDIVSOIKTQLOSDITSFRNSAIBALNRFIOKYSVMQRLDDTSRK"

BASE COUNT 343 a 184 c 205 g 282 t
 ORIGIN
 alignment_scores:
 Quality: 1641.00 Length: 325
 Ratio: 5.049 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 99.692

alignment_block:
US-08-699-716a-2 x A56795

Align seg 1/1 to: A56795 from: 1 to: 1014

```

197 ILAARGALATYRGLUGINASNPROGLNHISPHLEGLUASPLEUGLU 213
13 ATTCGACCCACGAAACCAACCAACATTTATTGAGATCTAGAAAA 62
213 SVALARGVALGLUGINLEUTHRGVHISGLYSERSEVALLEUGLU 230
63 AGTTAGGCTGGAACACTTACTGCTACTGCTCTCTGCTTTCAGTAA 112
230 EUVALGLINLEUVALYASPLYSASNILEASPILESERILEYSTYR 246
113 TGGTTCAGTTAGTCAAGATAAATAATAGTATTTCCATTAAATGAT 162
247 PROARGLYASPSERGLUVALPHEALASNAARYALLLETHRASPA 263
163 CCCAGAAAATTCGAGGTTTCCCAATAGATTAATCTGATGATAT 212
263 EGLULEULEULYSYRILEUVALATYRPHLEUPROGLUASPHRIEL 280
213 CGAATTCCTCAAGAAAATCCTAGCTTATTTCTACCCGAGATGCCAT 262
280 EULYSGLYGLYHISTYRASPASNGLINLEUGNASNGLYILEYAR 296
263 TTAAGGCGGTCAATTATGACACCACTGCAAAATGCGATCAACGAG 312
297 LYGGLUHEUGLUSERSERPROASNTHGLINTPGLUENALGALPH 313
313 AAGAGTCTTGATATCGCCGAATACCAATGGAATGCGCGCTT 362
313 EMETALAVALMETHISPHESERLEUTHRALASPARILEASPSAPI 330
363 CARGCGATATGATCTTCTTTAACCCGCGATGATCATGATATATA 412
330 LEULYSVALILEVALASPSERMETASNHSISGLYASPALAARG 346
413 TTTTGAAGATGATGTGATTCATGATCATGATGATGATGCGCGT 462
347 LYSLEUARGLUGLULEUVALAGLULEUTHRALAGLULEULYSI 363
463 AAGTGGGTGAAGATGATGATGATGATGATGATGATGATGATG 512
363 VALILEGLALAGLULEUVALASNLYSHISLEUSERSERGLYTHRI 380
513 AGTTATTCACACCCGAATTTAATAGCATCTGCTAGTAGTGCCAT 562
380 SNILEHASPLYSERILEASNLEUMETASPLYSASNLEUTYRGL 396
563 ATATCCATGATTAATCATTAATCTCATGATTAATAAATTAATAG 612
397 THRASPLUGLULEUPHELYSALASERLAGLUYRYSILEUGLU 413
613 ACAATATAAGATTTTAAAGCCAGCGAGATCAAAATTCGAGAA 662
413 SMEPRGGLTHTRHLEGLINVALASPLYSERGLYSLYSLEVALS 430
663 AATGCTCAACACCATTCAGTGGATGGAGGAGAGAAAAAATAGCT 712
430 ERILEYASPHLEUGLYSERGLUASNLYSARGTHRGYALALEUG 446
713 CGATAAAGACTTCTTGGAAGTGAATAAAGAACCGGCGCTTGCT 762
447 ASNLEULYSASNSEYRSETYRASNLYSASPASNANGLUSEU 463
763 AACTGTAAAACTACTCTTATATAAAGATTAATGATATATCTCA 812
463 SPHEALATHTRCHYSERASPLYSERARGPROLEUASNAPLEU 480

```

```

813 CTTGGCACCACTGCTGCGATAGTCCAGCGCGCTCAACGATCTGTTA 862
480 ERGLNYSYRTHRGINLEUSERASPILETHSERARGPHEASERALA 496
863 GCCAAAAACACTGCTGATATATACATCAGCTTTAATGAGCT 912
497 ILGLUALLEUASNARGPHELEGLINLYSTYRASPSEVALMETCL 513
913 ATTTAAGCACTGAACCGTTTCATTCAGAAAATGATGATGACAGC 962
513 GLEULEASPSPTHSERGLYLYS 521
963 TCTGCTGATGACACGCTGCTGTTAA 987

```

seq_name: gb_baz:AF167309

seq_documentation_block: 975 bp DNA

DEFINITION Yersinia pestis strain Pestoides F v antigen (icrv) gene, complete

ACCESSION AF167309

VERSION AF167309.1 GI:7578512

KEYWORDS Yersinia pestis.

SOURCE Yersinia pestis.

REFERENCE 1 (bases 1 to 975)

Autor Adams, D.M., Worsham, P.L., Hill, K.K., Klevytska, A.M., Jackson, P.J.,

Friedlander, A.M. and Keim, P.

Diversity in a variable-number tandem repeat from Yersinia pestis

J. Clin. Microbiol. 38 (4), 1516-1519 (2000)

MEDLINE 20211685

REFERENCE 2 (bases 1 to 975)

Hill, K.K. and Jackson, P.J.

Direct Submission

Submitted (08-JUN-1999) Bioscience Division, Los Alamos National

Laboratory, MS:888, Los Alamos, NM 87545, USA

FEATURES

Location/Qualifiers

1..975

/organism="Yersinia pestis"

/strain="Pestoides F"

/db_xref="taxon:632"

1..975

/gene="icrv"

1..975

/gene="icrv"

/codon_start=1

/transl_table=11

/product="v antigen"

/protein_id="AAF64076.1"

/db_xref="GI:7578513"

/translation="MIRAYEONPOHFIEDLEKRVVEOLTGSSVLELYVOLVDKNI

DISTKYDRKRSEVFNANVTDDIELKILATYLPEDAILKGGHYNOLONGKRYK

EFLESPPTOWELRAFNAVHRSLEADRIDDDIKVYDSNNHGDASKURELAE

TAELKITSVIOAEIKHLSSGGINIHDKSINLMDKNLYGTDEIRAKSEYRILEK

MPOTTIOVDSSEKIVSIKIDLGSENKRTGALGINKSYGNKDNNEISHFATTCSPK

SRPLNDVLSQTTOLSDITSRFNSAIEALNRFIOKYSVMQRLDDTR"

BASE COUNT 334 a 174 c 194 g 273 t

ORIGIN

alignment_scores:

Quality: 1631.00 Length: 323

Ratio: 5.050 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 99.690

alignment_block:
US-08-699-716a-2 x AF167309

Align seg 1/1 to: AF167309 from: 1 to: 975

196 MCTLEARGALATYRGLUGINASNPROGLNHISPHLEGLUASPLEUGLU 212

```

seq_name: gb_baz:AF167310
512 nargLeuLeuaspapThr 518
|||||
951 ACCTGCTGCTAGATGACAGC 969
|||||

seq_documentation_block:
LOCUS AF167310 981 bp DNA BCT 17-APR-2000
DEFINITION Yersinia pestis strain Angola V antigen (lcrV) gene, complete cds.
ACCESSION AF167310
VERSION AF167310.1 GI:7578514
KEYWORDS
Yersinia pestis.
SOURCE Yersinia pestis
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
REFERENCE 1 (bases 1 to 981)
AAalt,D.M., Worsham,P.L., Hill,K.K., Klevytzka,A.M., Jackson,P.J.,
Friedlander,A.M. and Kelm,P.
Diversity in a variable-number tandem repeat from Yersinia pestis
J. Clin. Microbiol. 38 (4), 1516-1519 (2000)
20211685
REFERENCE 2 (bases 1 to 981)
Hill,K.K. and Jackson,P.J.
Direct Submmission
Submitted (08-JUL-1999) Bioscience Division, Los Alamos National
Laboratory, MS-1888, Los Alamos, NM 87545, USA
FEATURES
source
location/Qualifiers
1..981
/organism="Yersinia pestis"
/strain="Angola"
/db_xref="taxon:632"
1..981
/gene="lcrV"
1..981
/gene="lcrV"
/codon_start=1
/transl_table=11
/product="V antigen"
/protein_id="AA64077.1"
/db_xref="GI:7578515"
/translation="MIRAYDQNPQHFIEDLENVQELTGHGSSVLELYVOLVKNKNI
DIISKYPRKDSVEFARVITDDIELKRLIAYFLPEDAILKGHYNOLONGIKRVK
EFPSSNQTQWELRAFVAVHFSILDRIDDILKTVDSNMNHRGARSRLRELAEL
TALKIKTVQIAELINKLSSGGINIHDKSINLMDKRLCYTDEIRKASAEYILLEK
MPQTTIOVDSSEKKIVISIKDFLSENKRTALGLKNSYSINKNDELSHFATYSQK
SRPLNDIVSQKTTQLSDITSIRFSALEALNRFQKDYDSVMORLLDITSGK"
BASE COUNT 334 a 175 c 196 g 276 t
ORIGIN
Alignment_scores:
Quality: 1628.00 Length: 336
Ratio: 5.009 Gaps: 0
Percent Similarity: 99.693 Percent Identity: 98.773
Alignment_block:
US-08-699-716A-2 x AF167310 ..
Align seg 1/1 to: AF167310 from: 1 to: 981
196 MetIleArgAlaTyrGlUGlnAsnProGlnHISpelleGluAspLeuG 212
|||||
1 ATGATTAGAGCCTACGAAACAAACCCACACACTTTATTAGAGATCTAGA 50
|||||
212 uTyValArgValGIuGlnLeuThrGIYHISglYserSerValleuGluG 229
|||||
51 AAATGTTAGGTCGGAACAACACTTACTGTCATGCTTCTTCAGTTTAAAGA 100
|||||
229 IuLeuValGIuLeuValIysAspIysAsnIleAspIleSerIleIysTyr 245
|||||

```

```

101 AATTGCTTCACTAGTCAAGATATAATATAGATTTCATTAAATAT 150
246 AspProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
151 GATCCAGAAAAGATCGAGGTTTGGCCATAGAGTAATACTGATGA 200
262 PileGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThr 279
201 TATCGAATGCTCAGAAAATCCCTAGCTTATTCTACCCAGAGATCCA 250
279 IeLeuLysGlyGlyHisTyrAspAsnGluLeuGlnAsnGlyIleLysArg 295
251 TTCTTAAAGCGGCTCATTATGACAACTGCAAAATGGCATTAAGCGA 300
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTyrGluLeuArg 312
301 GTAAGAAGTCTCTGAAATCATCCCGAATACACAATGGGAATGGCGGC 350
312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
351 GTTCATGGCAGTAATGCATTTCTTAAACCGCGATCGATCATGATG 400
329 spIleLeuLysValIleValAspSerMetAsnHisGlyAspAlaArg 345
401 ATATTGAAAAGTATTGTTGATTCATCAATGAAATCATGATGATCCCT 450
346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIle 362
451 AGCAAGTGGCGTGAAGAATAGCTGAGGTTACCCCGAATTAAGAATTA 500
362 rSerValIleGlnAlaGluIleAsnLysHisLeuSerSerGlyThr 379
501 TTCAGTATTCAACCGCAATTAATTAAGCATCTGCTAGTGGGACCA 550
379 IeAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
551 TAAATATCCATGATTAATCCATTAATCATGATGATAAAATTTATAGT 600
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeu 412
601 TATACAGATGAAGATTTTAAAGCCGACGAGTCAAAATTCGCA 650
412 uLysMetProGlnThrThrIleGlnValAspGlySerGluLysIle 429
651 GAAATGCGCTCAAAACCACTTCAGTGTGATGGGACGAGAAAAAATAG 700
429 aLseIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
701 TCTCGATTAAGGACTTCTTGAAGTGAATTAAGAAGACGGGGCGGTG 750
446 GlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeu 462
751 GGTATCTGAAAACCTACTACTATATAAAGATATAATATGATATTC 800
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeu 479
801 TCACATTCACACACACAGCTCGATTAAGTCCAGCGCTCAACGACTGG 850
479 aLseGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
851 TTAGCCAAAACAACTCAGCTGCTGATTAATACATCAGTTTAAATCA 900
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG 512
901 GCTATGAGACATGACCGCTTTCATTCAGAAATATGATCAGTATGCA 950
512 nArgLeuLeuAspAspThrSerGlyLys 521
951 ACGTCTGCTAGATGACACGCTGATAA 978
seq_name: gb_ba1:YPPPIVANT
seq_documentation_block:

```

```

LOCUS YPPPIVANT 981 bp DNA BCT 13-MAR-1997
DEFINITION Y.pseudotuberculosis V antigen gene.
ACCESSION X96802
VERSION X96802.1 GI:1405834
KEYWORDS V antigen.
SOURCE Yersinia pseudotuberculosis.
ORGANISM Yersinia pseudotuberculosis.
REFERENCE 1 (bases 1 to 981)
AUTHORS Rogenkamp, A., Geiger, A.M., Leitritz, L., Kessler, A. and
Heesemann, J.
TITLE Passive immunity to infection with Yersinia spp. mediated by
anti-recombinant V antigen is dependent on polymorphism of V
antigen
JOURNAL Infect. Immun. 65 (2), 446-451 (1997)
MEDLINE 97162308
REFERENCE 2 (bases 1 to 981)
AUTHORS Rogenkamp, A.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-1996) A. Rogenkamp, Hygiene and Microbiology,
University of Wuerzburg, Josef-Schneider Strasse 2, 97080
Wuerzburg, FRG
FEATURES
source location/Qualifiers
1..981
/organism="Yersinia pseudotuberculosis"
/sub_species="type I"
/db_xref="taxon:633"
1..981
/codon_start=1
/transl_table=11
/product="V antigen"
/protein_id="CA65594.1"
/db_xref="GI:1405835"
/db_xref="EMBL:U00001"
/translation="MIRAEONPQHFIDLEKRYVEQITGSGSYLEEVLYVDKNI
DISIKYDRKDSYEVANRVITDDLEKILATYLPETAILKGGHYDQQLQNGIKRYK
EFLESPNTQWELRAFMAVIFHSLADRIDDDILKVIYDSNMHGDANSKLRBLAEL
TAEIKYIVIOAEIKHLSGSGTINIHRSINLMDKNYGYTDEIFKASAEYKILEK
MQGTIOGETEKIKVSKNLESEKRRKGALGNLKSYSYKNKNELSHFATVCSDK
SRPLNDIVSQKTLTQSDTSRFSNLEIENFIDKIDYSVMRLDDTISGK"
BASE COUNT 341 a 174 c 191 g 275 t
ORIGIN
alignment_scores:
Quality: 1590.00 Length: 326
Ratio: 4.938 Gaps: 0
Percent Similarity: 98.773 Percent Identity: 96.626
alignment_block:
us-08-699-716a-2 x YPPPIVANT
Align seg 1/1 to: YPPPIVANT from: 1 to: 981
196 MetIleArgAlaTyrGluGlnAsnProGlnHisPheIleGluLysLeuG 212
1 ATGATTAGAGGCTACGAAACAAACCAACATTTATTGAGATCTAGA 50
51 AAAAGTAGGCTGCAACACTGCTACTGCTCTTCAGTTTGAAG 100
212 uLysValArgValGluGlnLeuThrGlyHisGlySerSerValLeuGlu 229
51 AAAAGTAGGCTGCAACACTGCTACTGCTCTTCAGTTTGAAG 100
229 ILeuLeuValGlnLeuValLysAspLysAsnLleAspIleSerIleLysTyr 245
101 AATTGCTTCACTAGTCAAGATATAATATAGATTTCATTAAATAT 150
246 AspProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
151 GATCCAGAAAAGATTCGAGGTTTGGCCATAGAGTAATACTGATGA 200
262 PileGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThr 279

```

201 TATCGAATGCTCAAGAAAATCCTAGCTTATTTCTACCCGAGATGCCA 250
 279 ldeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyLeuSarg 295
 251 TCTTAAGCGCGTCATTTATGACACCACTGCAAAATGGCATCAAGCA 300
 296 ValLysGlnPheLeuGlnUserSerProAsnThrGlnTyrGlnLeuArg 312
 301 GTAAAGAGTTCCTTGAATCATCGCCGATACACAATGGGAATGGCGGC 350
 312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
 351 GTTCATGCGAGTAATACATTTCTCTTTAACCGCCGTCCTATCGAGATG 400
 329 spileuLysValIleValAspSerMetAsnHisIleGlyAspAlaArg 345
 401 AATTTTGAAGTGAATGTGTGATTCATCAATCATCATGTGATGCCCT 450
 346 SerLysLeuArgGlnGlnLeuAlaGlnLeuThrAlaGlnLeuLysIleTy 362
 451 ACCAAGTTCGCGAAGAATTACTGAGCTTACCGCCGATTAAGATTAA 500
 362 rSerValIleGlnAlaGlnIleAsnLysHisLeuSerSerGlyThrI 379
 501 TTCAGTATTCAGCCGAAATTAATAGCATGCTGTAGTGGCCACCA 550
 379 lAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
 551 TAAATATCATGATTAATCAATTAATCATGATGATTAATTAATTTATG 600
 396 TyrThrAspGlnGlnIlePheLysAlaSerAlaGlnTyrLysIleLeu 412
 601 TATACAGATGAGAGATTTTAAAGCCAGTGCAGACTACAAATTCGCA 650
 412 uLysMetProGlnThrThrIleGlnValAspGlySerGlnLysIleY 429
 651 GAAATATGCTCAACACCATTACAGAGAGTGTAGAGCGAAGAAAAATAG 700
 429 aSerIleLysAspPheLeuGlySerGlnAsnLysArgThrGlyAlaLeu 445
 701 TCTCGATTAAGACATTTCTTGAAGTGAAGAAAAAGAACCGGCGCTTG 750
 446 GlnAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGlnLeuSe 462
 751 GGTAATCTGAAGACATCACTCTTATTAATAAGATTAATGAATATC 800
 462 HisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
 801 TCACCTTGGCCACCCTGCTCGATTAAGTCCAGCCGCTCAACGACTTGG 850
 479 aISerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
 851 TTAGCCAAAACAACTCACTGCTGATTAATTAACACAGCTTTAATTA 900
 496 AlaIleGlnAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG 512
 901 GGTATGAGCACTGAACGCTTTCATTCAGAAATATAGATTCAGTATGCA 950
 512 nArgLeuLeuAspAspThrSerGlyLys 521
 951 ACCTCTGCTAGATGACACCTGTGTAAA 978
 seq_name: qb_bal:YEPICRGVHP
 seq_documentation_block:
 LOCUS YEPICRGVHP 2201 bp DNA BCT 26-Apr-1993
 DEFINITION Yersinia pseudotuberculosis V-antigen (LcrG, LcrV, LcrH) genes,
 complete cds.
 ACCESSION M57893
 VERSION M57893.1 GI:155456
 KEYWORDS V-antigen,
 SOURCE Yersinia pseudotuberculosis (strain YPIII (PIBI) DNA.
 ORGANISM Yersinia pseudotuberculosis

REFERENCE
 1 (bases 1 to 2201)
 AUTHORS
 Bergman,T., Hakanson,S., Forsberg,A., Norlander,L., Macellaro,A.,
 Beckman,A., Boelln,T. and Wolf-Watz,H.
 TITLE
 Analysis of the V antigen lcrGH-yopBD operon of Yersinia
 pseudotuberculosis: Evidence for a regulatory role of LcrH and LcrV
 JOURNAL
 Bacteriol. 173, 1607-1616 (1991)
 MEDLINE
 91154114
 FEATURES
 source
 location/Qualifiers
 1..2201
 /organism="Yersinia pseudotuberculosis"
 /strain="YPIII (PIBI)"
 /db_xref="taxon:633"
 219..224
 /gene="lcr operon"
 219..258
 /gene="lcr operon"
 241..246
 /gene="lcr operon"
 252..258
 /gene="lcr operon"
 264..551
 /gene="lcrG"
 264..551
 /gene="lcrH"
 /codon_start=1
 /transl_table=1
 /evidence=experimental
 /product="V-antigen"
 /protein_id="AA27644.1"
 /db_xref="GI:155457"
 /translation="MKSSHFDKRLKQALAIADSDRAKILQEMCDIGITPNAV
 MKTFAKSAEELPAERLEIDKRLKROPQHPYDGGKPKPTMKGQIT"
 553..1533
 /gene="lcrV"
 553..1533
 /gene="lcrV"
 /codon_start=1
 /transl_table=1
 /evidence=experimental
 /product="V-antigen"
 /protein_id="AA27645.1"
 /db_xref="GI:155458"
 /translation="NIRAYBONQHRIEDLEKRVQVLGHGSSYLEVQLYKDKNI
 DIKIKYDKRSEVFANRVITDDIELKILAFLEPDAILKGHDNOLQNGIKRVK
 EFLSESPNTOEWELRAPMAVHESLFDRIIDDLIKYIVDSNMHGDARSLRELBEL
 TALKIYSYQAEINRHLSSGGFINIDKSLINDKILGYTDEIFKSAEYKILEK
 MPQTTIOEGTEKIKYSIKNFLESEKKRKGALGNLDSYKNKDNELSHFATGCDK
 SRPLNDLVSGKTQLSDITSRFSATLEALNRFQKYDSVMGRLLDPTSGK"
 1546..2052
 /gene="lcrH"
 1546..2052
 /gene="lcrH"
 /codon_start=1
 /transl_table=1
 /evidence=experimental
 /product="V-antigen"
 /protein_id="AA27646.1"
 /db_xref="GI:155459"
 /translation="MQERTPDQYOLAMESFLKGGGTIAMNEISSDLYSLAF
 NOYOSKYEPAHVKVFOALCYLDHRYDFRFLGACACQAMQOYIOLAHSSYSGAIINIK
 ERFPPHAACCLLQKELDAESGLFLAOLADKPEFELSTRVSSMLPAIKREK
 EHECVNP"
 BASE COUNT 714 a 450 c 485 g 552 t
 ORIGIN
 alignment_scores:
 Quality: 1590.00 Length: 326
 Ratio: 4.938 Gaps: 0
 Percent Similarity: 98.773 Percent Identity: 96.626

alignment_block:
US-08-699-716a-2 x YEPLCRGHP ..

Align seg 1/1 to: YEPLCRGHP from: 1 to: 2201

```

196 MetIleArgAlaTyrGluInaSerProGlnHisPheIleGluAspLeuGlu 212
|||||
553 ATGATTAGACCTACGACAAACCCACAACTTTATTTAGAGATCTAGA 602
212 ulysValArgValGluGlnLeuThrGlyHisGlySerSerValLeuGluG 229
|||||
603 AAAAGTTAGGCTGGAACAACTTACTGTCTAGTCTTCTAGTTTAGAG 652
229 IuLeuValGlnLeuValLysAspLysAsnIleAspIleSerIleLysTyr 245
|||||
653 AATGTGCTCAGTAGTCAAAGATAAAATATAGTATTTCCATTAAATAT 702
246 AspProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
|||||
703 GATCCCGAAGAAAGATTGGAGGTTTGGCCAAATAGACTAATTTCTGAT 732
262 pIleGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThrI 279
|||||
753 TATCGAATTCCTCAGAAATCTAGCTTATTTCTACCCGAGATGCCA 802
279 IeLeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
|||||
803 TTCTTAAGCGGTCATTATGACACCACTGCAAAATGGCATCAGACGA 852
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTrpGluLeuArgAl 312
|||||
853 GTTAAAGAGTTCCTTGAATCATCGCCAAATACAAAGGGAATGGCGGC 902
312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAspA 329
|||||
903 GTTCATGCGCAGTATACATTTCTTTAACCGCGCATGATGATGATG 952
329 spIleLeuLysValIleValAspSerMetAsnHisGlyAspAlaArg 345
|||||
953 ATATTTTGAAGATGATTGATTCATCATCATCATGATGATGATGATG 1002
346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleT 362
|||||
1003 AGCAAGTTGCGTGAAGATTTAGCTGAGCTTACCGCCAAATTAATAAT 1052
362 rSerValIleGlnAlaGluIleAsnLysHisLysSerSerSerGlyThrI 379
|||||
1053 TTCAGTTATTCAGAGCTGAATTAATTAAGCATCTGTGCGAGTGGCCCA 1102
379 IeAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
|||||
1103 TAAATATCCATGATTAATCAATTAATCTCATGATTAATAATTAATAG 1152
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuG 412
|||||
1153 TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAATTCGGA 1202
412 ulysMetProGlnThrThrIleGlnValAspGlySerGluLysLysIle 429
|||||
1203 GAAATAGCTCAAAACCAACCATTCAGGAAGGTGAGACCGAAAAAATAG 1252
429 aIleSerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
|||||
1253 TCTCGATTAAGAACTTCTTGAAGAGTGAAGAAAAAGAACCGGCGCTTG 1302
446 GlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnGlnLeuLeu 462
|||||
1303 GGTAAATCTGAAGACATCATCTTATAAATAAGATATATATGAATATTC 1352
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
|||||
1353 TCACCTTGGCACCACTGCTCGGATAGTCCAGGCGCTCAAGACACTGG 1402

```

```

479 aIleSerIleLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
|||||
1403 TTAGCCAAAAAACAACTCACTGTCTGATATATACATACAGCTTTTATTTCA 1452
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetI 512
|||||
1453 GCTATTGAGCACTGAACCGTTTCATTTCACAAATATGATTCAGTATGCA 1502
512 nArgLeuLeuAspAspThrSerGlyLys 521
|||||
1503 ACCTCTGCTAGATGACACGCTGTGTAAA 1530
seq_name: gp_baz:AF102990
seq_documentation_block:
LOCUS AF102990 69673 bp DNA circular BCT 18-MAY-1999
DEFINITION Yersinia enterocolitica plasmid pYE227, complete sequence.
ACCESSION AF102990 AF054979 AF054980 AF054981 AF080156 Z69926 L06216
AF080155 AF022645 AF050104 AF054977 AF033863 U02499 U08019 U21297
M22781 U08222 Z18539 U94827
AF102990.1 GI:4324323
VERSION
KEYWORDS
SOURCE
ORGANISM
Yersinia enterocolitica.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
1 (bases 32056 to 32871)
Cornelis, G., Sluifers, C., de Rouvoit, C.L. and Michiels, T.
Homology between virF, the transcriptional activator of the
Yersinia virulence regulon, and AraC, the Escherichia coli
arabinose operon regulator
J. Bacteriol. 171 (1), 254-262 (1989)
89123026
2 (bases 15074 to 15994)
Hakansson, S., Bergman, T., Vanooteghem, J.C., Cornelis, G. and
Wolf-Watz, H.
YopB and YopD constitute a novel class of Yersinia Yop proteins
Interf. Immun. 61 (1), 71-80 (1993)
93114907
3 (bases 45325 to 45717)
Wattiau, P. and Cornelis, G.R.
SydE, a chaparrone-like protein of Yersinia enterocolitica involved
in Ome secretion of YopE
Mol. Microbiol. 8 (1), 123-131 (1993)
93268087
4 (bases 23950 to 25269)
Woestyn, S., Allaoui, A., Wattiau, P. and Cornelis, G.R.
YscN, the putative energizer of the Yersinia Yop secretion
machinery
J. Bacteriol. 176 (6), 1561-1569 (1994)
94179088
5 (bases 29897 to 30961)
Allaoui, A., Woestyn, S., Sluifers, C. and Cornelis, G.R.
YscD, a Yersinia enterocolitica inner membrane protein involved in
Yop secretion
J. Bacteriol. 176 (15), 4534-4542 (1994)
94321323
6 (bases 43292 to 43717)
Wattiau, P., Bernier, B., Deslee, P., Michiels, T. and Cornelis, G.R.
Individual chaparrones required for Yop secretion by Yersinia
Proc. Natl. Acad. Sci. U.S.A. 91 (22), 10493-10497 (1994)
95024141
7 (bases 31537 to 31932)
Allaoui, A., Scheen, R., Lambert de Rouvoit, C. and Cornelis, G.R.
VirG, a Yersinia enterocolitica lipoprotein involved in Ca2+
dependency, is related to exsB of Pseudomonas aeruginosa
J. Bacteriol. 177 (15), 4230-4237 (1995)
95362644
8 (bases 53135 to 53497)
Stainier, I., Irlarte, M. and Cornelis, G.R.
YscM1 and YscM2, two Yersinia enterocolitica proteins causing
downregulation of yop transcription
Mol. Microbiol. 26 (4), 833-843 (1997)
JOURNAL

```



```

201 TATGGAATGCTCAGAAAATCCTAGCTTATTCTTACCCGAGATGCCA 250
279 leuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
251 TTCTTAAAGCGGCTATTATGACACCACTGCAAAATGGCATACACGA 300
296 ValLysGlnPheLeuGlnSerProAsnThrGlnTrpGluLeuArgAl 312
301 GTAAAGAGTTCCTTGATCATCGCCAAATACACATGGGAATGGCGGC 350
312 aspMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
351 GTTCATGCGAGTATACATTTCTTACCGCGCATGATCATGATGATG 400
329 spLleLeuLysValIleValAspSerMetAsnHisHisGlyAspAlaArg 345
401 ATATTGAAAGATGATGTTGATTCATCAATCAATCATGGTATGCCCGT 450
346 SerLysLeuArgGlnGluLeuAlaGluLeuThrAlaGluLeuLysIleTy 362
451 AGCAAGTGGGTGAGAAATTAAGTCTTACCGCGCATTAAGATTTA 500
362 rSerValIleGlnAlaGluIleAsnLysHisLeuSerSerSerGlyThrI 379
501 TTCAGTATTCAACGCCAAATTAATAGCATCTGTAGTGGTGACACA 550
379 leaAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
551 TAAATATCCATGATTAATCAATATCATGATGATTAATAATTAATATG 600
396 TyrThrAspGlnGluIlePheLysAlaSerAlaGluTyrIleLysIleGlu 412
601 TATACATGAGAGATTTTAAAGCCAGTGCAGATTCATCAAAATCTCGA 650
412 uLysMetProGlnThrIleGlnValAspGlySerGluLysLysIleVal 429
651 GAAATGCTTCAACCAACCATTCAGAGAGTGAAGCCGAAACCAAAATAG 700
429 alSerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
701 TCTCGATTAAGAACTTCTTGAAAGTGAGAAAAGAAACCGGCGCTTG 750
446 GlnValAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe 462
751 GGTATCTGAAAGACTACTACTTATATTAATTAATTAATTAATTAATTC 800
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuVal 479
801 TCACTTGGCCACACCTGCTCGATTAAGTCCAGGCGCTCAACGACTTGG 850
479 alSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
851 TTACCCAAAAAACAACCTGCTGATTAATTAATTAATTAATTAATTAATCA 900
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG1 512
901 GCTATTGAGACATGACCGTTTCAATTCAGAAATATGATTCAGTATGCA 950
512 narGluLeuAspAspThrSerGlyLys 521
951 ACGTCTGCTAGATGACAGCTCTGGTAA 978

```

seq_name: gb_ba1.YEPLCRGVP

seq_documentation_block: 2201 bp DNA BCT 26-APR-1993
 LOCUS YEPLCRGVP
 DEFINITION Yersinia pseudotuberculosis V-antigen (LcrG, LcrV, LcrH) genes,
 complete cds.
 ACCESSION M57893
 VERSION M57893.1 GI:155456
 KEYWORDS V-antigen.
 SOURCE Yersinia pseudotuberculosis (strain YPIII (PIB1) DNA.
 ORGANISM Yersinia pseudotuberculosis

```

REFERENCE 1 (bases 1 to 2201)
AUTHORS Bergman,T., Hakansson,S., Forsberg,A., Norlander,L., Macellaro,A.,
Baekman,A., Boelin,I. and Wolf-Watz,H.
TITLE Analysis of the V antigen lcrGVH-yopD Operon of Yersinia
JOURNAL pseudotuberculosis: Evidence for a regulatory role of LcrH and LcrV
MEDLINE J. Bacteriol. 173, 1607-1616 (1991)
FEATURES
source location/Qualifiers
1..2201 /organism="Yersinia pseudotuberculosis"
/strain="YPIII(PIB1)"
/db_xref="taxon:633"
-35_signal 219..224
gene 219..258
-10_signal 241..246
RBS 252..258
gene 264..351
CDS 264..551
/gene="lcrG"
/codon_start=1
/transl_table=1
/evidence=experimental
/product="V-antigen"
/protein_id="AAA27644.1"
/db_xref="GI:155457"
/translation="MKSSHPEYDYLTKQAEILADSPRAKILQEMADIGLPEAV
KTIKFGRAEELKPAERELDEIKRERQPHDPGRKPRKPTMARQIIT"
553..1533
/gene="lcrV"
553..1533
CDS 553..1533
/gene="lcrV"
/codon_start=1
/transl_table=1
/evidence=experimental
/product="V-antigen"
/protein_id="AAA27645.1"
/db_xref="GI:155458"
/translation="MIRAEQNPQHFIEDLEKVRVQGLGHSSTVLELYVQYKDKNI
DISIKYDKDSVFNARVTDIDIELKILATFPEDAILGSHYDQNGIRKVK
EFLSEPTQWELARFMAVIRFSLADRIDDLIVDSNMHEDASKLELAEI
TAEKIVSVIOAEINKHLSSGGINIHDKSLMDKNLYGDEIEFASABYKILEK
MNOTIOEGETEKKIVSIKNFLESKKRTGALGNKDSYKRNKNNELSHFATCSDK
SRPLDIVSOKTQQLSDITSRNSAIEALNRIQKYDSVMQRLDDTSK"
1546..2052
/gene="lcrH"
1546..2052
CDS 1546..2052
/gene="lcrH"
/codon_start=1
/transl_table=1
/evidence=experimental
/product="V-antigen"
/protein_id="AAA27646.1"
/db_xref="GI:155459"
/translation="MOERTDQEOVLAMESFLKGGGTIAMLNEISDTLEQYSLAF
NOYSGKYEDAKHYOALCVLDHYDSRFLGACROAMGOVDLAIHSGYGAINDIK
EPRFPNHAECILQGELEAEASGLFLAOELLTADPREFELSTRVSSMLEAIKLEKM
EHECVDP"
BASE COUNT 714 a 450 c 485 g 552 t
ORIGIN
alignment_scores:
Quality: 1590.00 Length: 326
Ratio: 4.938 Gaps: 0
Percent Similarity: 98.773 Percent Identity: 96.626

```

alignment_block:
US-08-699-716a-2 x YEPICRGVHP

Align seg 1/1 to: YEPICRGVHP from: 1 to: 2201

```

196 MetIeaTgAlaTyrGluGlnAsnProGlnHisPheIleGluAspLeuG1 212
553 ATGATTAAGGCTACGACAAACCAACCAACATTTTATGAGACTGACA 602
212 uLysValArgValGluGlnLeuThrGlyHisGlySerValLeuGluG 229
603 AAAAGTGGGTGGAACAACCTACTGCTCATGCTTCTTCAGTTTAAAG 652
229 LuLeuValGlnLeuValLysAspLysAsnIleAspIleSerIleLysTyr 245
653 AATTGCTTCACTTACTCAAGATTAATAATATGATTTTCCATTAAAT 702
246 AspProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
703 GATCCCAAGAAAAGATTCGAGGTTTTCGCAATAGAGTAATCTGATCA 752
262 PLeuGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThr 279
753 TATCAATATGCTCAAGAAATCTCATATTCTTACCCGAGATGCCA 802
279 LeLeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
803 TTCTTAAAGGGGGGTATTATGACAAACCACTGCAAAATGGCATCAAGCA 852
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTrpGluLeuArgAl 312
853 GTAAAGAAGTCTCTGATTCATCGCGCAATACAAATGGGATTCGGCGC 902
312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
903 GTTCATGCGACGATATACATTTCTTTAAACCCGATCGTATCATGATG 952
329 spIleLeuLysValIleValAspSerMetAsnHisHisGlyAspAlaArg 345
953 ATATTATTAAGTGAATGTGTGATTCATGATCAATGAAATCATGATGAT 1002
346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyr 362
1003 AGCAAGTGGCGGAGAAATGAGCTTACCGCGCAATTTAAATTTTA 1052
362 rSerValIleGlnAlaGluIleAsnLysHisLeuSerSerSerGlyThr 379
1053 TTCAGTTATTCAGCTGAATTAATTAACATCTGTCGAGTGGTGACACA 1102
379 leAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
1103 TAAATATCAATGATTAATCAATTAATCTCATGATTAATAATTAATGAT 1152
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuG1 412
1153 TATACAGATGAAGAGATTTTAAACCAACCGAGATACAAAATCTCGCA 1202
412 uLysMetProGlnThrThrIleGlnValAspGlySerGluLysIleVal 429
1203 GAAATATGCTCAAAACACCATTCAGGAAGTGAACGCAAAAAAAATAG 1252
429 aLserIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
1253 TCTGCATTAAGAACTTCTTGAAGTGAAGAAAAAGAAACCGCGCGCTGG 1302
446 GlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe 462
1303 GGTATATCGAAGACTCATCTTATATTAAGATTAATTAATGATTTATC 1352
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
1353 TCACITTTCCACCACTGCTCGGATAGTCCAGCGCGCTCAACGACTTGG 1402

```

```

479 aLserGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
1403 TTAGCAAAAAAACCACTAGCTGCTGATTAATTAATCACTTTTAATCA 1452
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG1 512
1453 GGTATTGAAGCACTGACACCGTTTCTCATTCGAAATATGATTCAGTATCA 1502
512 nArgLeuLeuAspAspThrSerGlyLys 521
1503 ACGTCTGCTAGATGACACGTCGTGTA 1530
seq_name: gb_baz:AF102990
seq_documentation_block:
LOCUS AF102990 69673 bp DNA circular BCT 18-MAY-1999
DEFINITION Yersinia enterocolitica plasmid pYez27, complete sequence.
ACCESSION AF102990 AF054978 AF054979 AF054980 AF054981 AF080156 269926 106216
AF080155 AF022645 AF050104 AF054977 AF033863 U02499 U08019 U21297
M22781 U08222 Z18539 U94827
AF102990.1 GI:4324323
VERSION
KEYWORDS
SOURCE
ORGANISM
Yersinia enterocolitica.
Yersinia enterocolitica
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
1 (bases 32056 to 32871)
Cornelis,G., Sulters,C., de Rouvoit,C.L. and Michiels,T.
Homology between yJrf, the transcriptional activator of the
Yersinia virulence regulon, and ArcC, the Escherichia coli
arabinose operon regulator
J. Bacteriol. 171 (1), 254-262 (1989)
89123026
2 (bases 15074 to 15994)
Hakansson,S., Bergman,T., Vanooteghem,J.C., Cornelis,G. and
Wolf-Watz,H.
YopB and YopD constitute a novel class of Yersinia Yop proteins
infect. Immun. 61 (1), 71-80 (1993)
93114907
3 (bases 45325 to 45717)
Wattiau,P. and Cornelis,G.R.
SYCE, a chaperone-like protein of Yersinia enterocolitica involved
in the secretion of YopE
Mol. Microbiol. 8 (1), 123-131 (1993)
93268087
4 (bases 23950 to 25269)
Moestyn,S., Allaoui,A., Wattiau,P. and Cornelis,G.R.
YecN, the putative energizer of the Yersinia Yop secretion
machinery
J. Bacteriol. 176 (6), 1561-1569 (1994)
94179088
5 (bases 29897 to 30961)
Allaoui,A., Moestyn,S., Sulters,C. and Cornelis,G.R.
Yacu, a Yersinia enterocolitica inner membrane protein involved in
Yop secretion
J. Bacteriol. 176 (15), 4534-4542 (1994)
94321323
6 (bases 43292 to 43717)
Wattiau,P., Bernier,B., Deslee,P., Michiels,T. and Cornelis,G.R.
Individual chaperones required for Yop secretion by Yersinia
Proc. Natl. Acad. Sci. U.S.A. 91 (22), 10493-10497 (1994)
95024141
7 (bases 31537 to 31932)
Allaoui,A., Scheen,R., Lambert de Rouvoit,C. and Cornelis,G.R.
Virg, a Yersinia enterocolitica lipoprotein involved in Ca2+
dependency, is related to exsB of Pseudomonas aeruginosa
J. Bacteriol. 177 (15), 4230-4237 (1995)
95362644
8 (bases 53135 to 53497)
Stainier,I., Irlarte,M. and Cornelis,G.R.
YscM and YscM2, two Yersinia enterocolitica proteins causing
downregulation of yop transcription
Mol. Microbiol. 26 (4), 833-843 (1997)
JOURNAL

```

MEDLINE
REFERENCE
AUTHORS
TITLE
98087284
(bases 18681 to 18968)
Boyd,A.P., Sory,M.P., Iriarte,M. and Cornelis,G.R.
Heparin interferes with translocation of Yop proteins into Hela cells and binds to LcrG, a regulatory component of the Yersinia Yop apparatus
Mol. Microbiol. 27 (2), 425-436 (1998)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
98143428
(bases 22612 to 22890)
Iriarte,M., Sory,M.P., Boland,A., Boyd,A.P., Mills,S.D., Lambermont,I. and Cornelis,G.R.
Tyea, a protein involved in control of Yop release and in translocation of Yersinia Yop effectors
EMBO J. 17 (7), 1907-1918 (1998)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
98190073
(bases 1 to 69673)
Iriarte,M., Lambermont,I., Kerbouch,C. and Cornelis,G.R.
Detailed genetic map of the pYve27 plasmid of Yersinia enterocolitica serotype O:9
Unpublished
12 (bases 1 to 69673)
Iriarte,M., Lambermont,I., Kerbouch,C. and Cornelis,G.R.
Direct Submission
Submitted (30-OCT-1998) Microbial Pathogenesis Unit, Christian de Duve Institute of Cellular Pathology and Faculte de Medecine, Universite Catholique de Louvain, Av. Hippocrate 74.49, Brussels 1200, Belgium

COMMENT
On May 18, 1999 this sequence version replaced g1:3411136
g1:3411135 g1:3411139 g1:3411137 g1:3603360 g1:1524380 g1:155536
g1:3603357 g1:2738986 g1:2944089 g1:3411138 g1:3108216 g1:437201
g1:497215 g1:885971 g1:155526 g1:483322 g1:48593 g1:2735421.
Location/Qualifiers
1..69673
/organism="Yersinia enterocolitica"
/strain="W22703"
/serotype="O:9"
/db_xref="taxon:630"
/plasmid="pYve27"
53..203
complement(889..1434)
/note="ORF181"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AADI6801.1"
/db_xref="GI:4324324"
/translation="MRVVIQGYGIGFKKNGLDLSMPAIOCTGLYALSDQHDYLCA
HEDSAFGLQONQDIRFTFLNGIRMSLRATVFGDGQSYRCSPASHIGNEIYN
FKMSGALAOYNOYSGVPEFNFHYQKGCITGQNPDRFLGSPQAMNAKQRI
KLRPSEYSLTHAKMDISRLY"
1616..2083
/note="ORF155"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AADI6802.1"
/db_xref="GI:4324325"
/translation="MINTFTLELQIASHFGDKLSODEYICELILNDRVYIMLRA
DEILNLTLLGLFSGPEARSTAFCCYSINLNDGPFANSEELGLAFKHL
SLGELNVENVSEIANFYDMLSLVSLPAETQOELPHTOSTOSTOSTOSVKMG"
2091..4280
/gene="yopO"
2091..4280
/gene="yopO"
/codon_start=1
/transl_table=11
/product="protein kinase YopO"
/protein_id="AADI6803.1"
/db_xref="GI:4324326"
/translation="MKIMGMPSPISLAKAHERISCHWONPVGELNIGKRRRIIDNO
VRLNPHSGFSLREGVGRIFSGKMFNSIARMLTTLHAQQTOSTOESDIPNLS
NLFGAPTEPLSGAPDKPLSGAPDEKGNVAVDTDLFAEGESHISIEIKRDLVAK
IERSIAGEHLFELVAKHYKHYKAGHPNLANVHGAAYVPYGRKREALLMDEVDMR

CSDTLRLSDSKCKKINSSEAYGTIKETIAHRLDVTNHLAKAGIVNDRKPNVFD
RASGEVYVLDLHRSGEOPKGFESKAPLGVGNIGASERKSVPLVYSLLGIE
GEEKDEIKPNGLRFTISEPAHVDENGYPHRRGIGVEATYRFTIDIGVADS
RPSDNEARLHETLSGTIDEBKAKIINDTLGEGSPSTDVRRTPPKRLRSOLR
THLSAATKQLDMGVSLDPLTMTVLTKAEREGGVDDXKDLDTQHSFSDIGSLVARHL
VKRGEDTKRSSAEVSPYHRSNFMLSIAEPSLQRLKLDQHSFSDIGSLVARHL
ETLLEVLVTLSPQGPVSEYSEFLNRLAEAVTLQOOLDLQOQOESAKADSLILN
RGSWADVAROSLOREDSTRPVKFGTEOYTAIHRQMAHAATITLOEVSFETDMRN
ETADSIPLILRRGRSLDEHVEOREKRLRELTITAEINRLREEM"
4485..4760
/note="ORF91A; unknown function"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AADI6804.1"
/db_xref="GI:4324327"
/translation="MSTIRKNEPCWTINNGLPVIGSELAINCKEFPVGINYSALYR
IMNLGVAVRVLPCSGFTHORLNVLLQGYVPEYKIRHAYVSLNDR"
4761..5627
/gene="yopP"
4761..5627
/gene="yopP"
/codon_start=1
/transl_table=11
/product="Yop effector YopP"
/protein_id="AADI6805.1"
/db_xref="GI:4324328"
/translation="WIGPISQNSFGSGSEKTRSLISNEELKNITIOLETDIADGWS
FHRKSRDIEVMPALVIOANKPYEMNLNFTSPQDLSIEIKNYIENGVSRRITIN
MOEGDIFSEVIDIKYKINGKSTLILEPVFNMGPAIIAISRTAIEROLDCDFSM
VMDGIRHSECEGIFSLAKRLYTERDSLLKIHNDNKGISIDSENPPLPHKRLDPL
PVTYFKHQGRKRLNBYLNTNPGGYVANKNENIIFNPNKNSIIDKELSVYHK
KRIAEYKILYK"
complement(6151..6900)
/gene="y1pA"
complement(6151..6900)
/gene="y1pA"
/codon_start=1
/transl_table=11
/product="lipoprotein y1pA"
/protein_id="AADI6806.1"
/db_xref="GI:4324329"
/translation="MEDDKKMKLLIATVAVLSVVLSCGAMSAIKRNLVYKTO
MSETIWMESPSSOKTYVLOIKNTSDKMKGLAKIRKANODKRYTTSSEDAHYVIOA
NVLKADKDLREABEFLSOGYGLALGALGAGITGYSNGASISVGLAGLGVY
ADAWVEDINTVAVTVQISEKTDPLQTDNVALLQGTISGYVQSTQGNHVOYTR
VVSANVKLKEEAOPLYEDOLANSIANIL"
complement(7370..7918)
/gene="yopQ"
complement(7370..7918)
/gene="yopQ"
/codon_start=1
/transl_table=11
/product="YopQ"
/protein_id="AADI6807.1"
/db_xref="GI:4324330"
/translation="MTKDAYMRLCTALREOSADPTIINTSEENNSYCATPAHLR
TVCGLVNRVGEIPLKSGSLTLELLEMAVGIIRLYEMOHVSDIDTFFKLPNNSD
FELVSVLDCDDIGYITGKDKSGKNGIETLDPNNSLIENDIRKYLIDENFRFCIML
IISKSELELSRESCDQKIMG"
8439..9407
/gene="yopT"
8439..9407
/gene="yopT"
/note="cytoxin"
/codon_start=1
/transl_table=11
/product="Yop effector YopT"
/protein_id="AADI6808.1"

alignment_scores: 1557.00 Length: 323
Quality: 4.896 Gaps: 0
Ratio:

Percent Similarity: -98.452 Percent Identity: 95.356

alignment_block:

us-08-699-716a-2 x AF102990/rev ..

Align seg 1/1 to reverse of: AF102990 from: 1 to: 69673

```
196 MetIleatgaIaTYrGlInAsnProGlnHsPheIleGluAspLeuG1 212
|||||
18679 ATGATTAGAGCTTCAGAACAAACCCACACATTATTATGAGACTGA 18630
|||||
212 uLysValrGValGluInleuThrGlyHsGlySerSerValleuG1 229
|||||
18629 AAAAGTTGGGTGGACACACTTACTGCTCATGGTCTTCAGTTTAGAG 18580
|||||
229 LuLeuValGlnLeuValLysAspLysAsnIleAspIleSerIleLysTyr 245
|||||
18579 AATTGGTTCAGTTAGTCAAGATAAAAGATGATATTCCATTAAATAT 18530
|||||
246 AsProAlGlyAspSerGluValPheAlAsnAlGValIleThrAspAs 262
|||||
18529 GATCCCAAAAAGATTCGGAAGTTTTCGCCATAGGTAATTACTGATGA 18480
|||||
262 PIIleGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThr 279
|||||
18479 TATCGAATTACTCAGAAATCTCTGCTTATTTCTACCCGAGATGCCA 18430
|||||
279 LeLeuLysGlyGlyHsTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
|||||
18429 TTCTTAAAGCGCGCATATATGACAACTGCAAAATGGCATCAAGCGA 18380
|||||
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTPGluLeuArgAl 312
|||||
18379 GTAAAGAGTTCTCTGAATCATCGCCGATACACATGGAGTTGCGGCG 18330
|||||
312 aPheMetAlaValMetHsPheSerLeuThrAlaAspArgIleAspAspA 329
|||||
18329 GTTCATGCGCAATATGCAATTTCTCTTTAACCCGCGATCGATCGATG 18280
|||||
329 sPIleLeuLysValIleValAspSerMetAsnHsIleGlyAspAlaArg 345
|||||
18279 AATTTTGAAGTGAATGATGTGATCATGAAATCATCATGATGATGCCGT 18230
|||||
346 SerLysLeuAlaGlyGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyr 362
|||||
18229 GGCAAGTTGCGTGAAGAAATAGCTAGCTTACCCCGAATTAAAGATTTA 18180
|||||
362 rSerValIleGlnAlaGluIleAsnLysHsIleuSerSerSerGlyThrI 379
|||||
18179 TTCAGTTATTCAAGCCGAATTAATTAAGCATCTGTCTAATAGTGACACCA 18130
|||||
379 LeAsnIleHsAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
|||||
18129 TAAATATCATCATGATTAATCAATTAATCATCATGATTAATAATTAATG 18080
|||||
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuG1 412
|||||
18079 TATACAGATGAGAGATTTTAAACCCAGCGAGATGACAAATTTCTCGA 18030
|||||
412 uLysMetProGlnThrIleGlnValAspLysSerGluLysLysIleVal 429
|||||
18029 GAAATGCTTCACACACCATTAAAGGAAGTGAGACCGAATAAATAATAG 17980
|||||
429 aLSerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
|||||
17979 TCTCGATAAGACACTTCTTGAAAGTGAGAAATAAAGAACCGGCGCTTG 17930
|||||
446 GlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe 462
|||||
17929 GGTAACTCGAAGATTCATCTATATATAAAGATAATGAATATATC 17880
|||||
462 rHsPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
|||||
```

```
17879 TCACCTTCCACACACCTGCTGCGATTAAGTCAGGCCGCTCAGACACTTGG 17830
479 aLSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
|||||
17829 TTAGCAAAAAACAACTCAGCTGCTGATATATACATCAGCTTTAATTCA 17780
|||||
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG1 512
|||||
17779 GCTATTGAAGCAGCTGAACCGTTTCATTCAGAAATATGACTCAGTATGCA 17730
|||||
512 nArgLeuLeuAspAspThr 518
|||||
17729 ACGTCTGCTAGATGACAG 17711
```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)


```

56 hrlvValGluProAlaArgIleThrLeuThrTyrIleGluValAlaPro 72
   |||
110 CTCTGTTGAACCGCCGATCCTTACATATATAGAGAGCGCTCA 159
   |||
73 IleThrIleMetAspAsnGlyAsnIleAspThrGluLeuLeuValGlyThr 89
   |||
160 ATTACATATTATGACAAATGGAACATCGATACGAATTCCTGTTGTCAC 209
   |||
89 rLeuThrLeuGlyGlyTyrIleThrGlyThrThrSerThrSerValAsp 106
   |||
210 GCTTACTCTTGGCGGTATATAAAGACAGAACCTACACATCTGTTAACT 259
   |||
106 heThrAspAlaAlaGlyAspPrometTyrLeuThrPheThrSerGlnAsp 122
   |||
260 TTACAGATGCCGCGGTATCCATGCTACTTAACATTTCTCTCGAGAT 309
   |||
123 GlyAsnAsnHISGlnPheThrThrIleValIleGlyAspSerArgAs 139
   |||
310 GGAATTAACCCACCATTCCTACAAAAGATGGCAGAGATTCCTAGACA 359
   |||
139 pPheAspIleSerProLysValAsnGlyLysLeuValGlyAspAsp 156
   |||
360 TTTGATATCTCTCCTAAGTAACGCTGAGAACCTGTGGGGGATGAGC 409
   |||
156 aIValLeuAlaThrGlySerGlnAspPheValArgSerIleGlySer 172
   |||
410 TCGTTTGGCTACGGGACCGACGAGATTTCTTGTTCCTCAATGGTTC 459
   |||
173 LysGlyGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrVal 189
   |||
460 AAGCGCGTAACCTGCACGAGTAATACACTGATGCTGAACCTAAC 509
   |||
189 rValSerAsnGln.....GluPheMetIleArgAlaTyrGluGln 203
   |||
510 CTTATCTAACCAAGATCCATCGAAGCTGTATAGAGCTTACGAAACAA 559
   |||
203 snProGlnHisPheIleGluAspLeuGluLysValArgValGluGlnLeu 219
   |||
560 ACCCAACAACTTTTATGAGATCTAGAAAAGTTAGGTGAAACAACCT 609
   |||
220 ThrGlyHisGlySerSerValLeuGluGluLeuValGluLeuValLys 236
   |||
610 ACTGTCATGCTTCTTCAGTTTACAGAAATGCTTCAGTTAGTCAACA 659
   |||
236 pLysAsnIleAspIleSerIleLysTyrAspProArgLysAspSerGly 253
   |||
660 TAAATATATAGATATTTCCATTAATATATGATCCAGAAAAGATTCG 709
   |||
253 aLPhaAlaAspArgValIleThrAspAspIleGluLeuLeuLysIle 269
   |||
710 TTTTGGCCATAGAGTAATTAAGATGATGAAATGCTCAGAAAATC 759
   |||
270 LeuAlaTyrPheLeuProGluAspThrIleLeuLysGlyIleHisTyr 286
   |||
760 CTAGCTTATTTCTACCCAGAGATGCAATCTTAAAGCGGCTCATATTA 809
   |||
286 pAsnGluLeuGlnAsnGlyIleLysArgValLysGluPheLeuGlnSer 303
   |||
810 CAACCAACTGCAAAATGCGATCAAGGAGTAAAGGTCCTGTAATCAT 859
   |||
303 exProAspThrGlnTrpGluLeuArgAlaPheMetAlaValMetHisPhe 319
   |||
860 CGCCGAATACAAATGGGAATGGCGGCTTATGCAATGCAATGCAATTC 909
   |||
320 SerLeuThrAlaAspArgIleAspAspAspIleLeuLysValIleVal 336
   |||
910 TCTTTAACCGCGGATCGTATCATGATGATATTTGAAAGTATGTTTGA 959
   |||
336 pSerMetAsnHisGlyAspAlaArgSerLysLeuArgGluGluLeu 353
   |||
960 TTCATATGATCATCATGATGATGCGGTGCAAGTTCGCTGAGAGATTTG 1009
   |||
353 IaGluLeuThrAlaGluLeuLysIleTyrSerValIleGlnAlaGluLe 369

```

```

|||||
1010 CTGAGCTTACCGCCGAATTAAGATTATTCAGTTATTCAGCCAAATT 1059
   |||
370 AsnLysHisLeuSerSerSerGlyThrIleAsnIleHisAspLysSerI 386
   |||
1060 AATAAGCATCTCTCTAGTATGTCACCATTAATATCCATGATGATCAT 1109
   |||
386 eaSnLeuMetAspLysAsnLeuTyrGlyTyrThrAspGluLeuIlePhe 403
   |||
1110 TATCTCATGATTAATAAATTTATATGTTATACAGATGAGAGATTTTGA 1159
   |||
403 yAlaSerAlaGluTyrLysIleLeuGluLysMetProGlnThrThrIle 419
   |||
1160 AAGCCAGCGCAGATACAAATTCGAGAAAATGCTCAACCCACCAT 1209
   |||
420 GlnValAspGlySerGlyLysLysIleValSerIleLysAspPheLeu 436
   |||
1210 CAGGTGATGGAGCGGAGAAAAAATAGTCTCGATTAAGGACTTCTTGG 1259
   |||
436 ySerGluAsnLysArgThrGlyAlaLeuGlyAsnLeuLysAsnSerTyr 453
   |||
1260 AAGTGAGATAAAGAACCGGCGCTGGTATCTGAAAACCTCATACT 1309
   |||
453 eTyrAsnLysAspAsnAsnGluLeuSerHisPheAlaThrThrCysSer 469
   |||
1310 CTTATTAATAAAGATATATATGATATCTCACTTGCACCACTGCTCG 1359
   |||
470 AspLysSerArgProLeuAsnAspLeuValSerGlnLysThrThrGln 486
   |||
1360 GATAGTCCAGGCGCCCTCAGACACTTGTGTAACCAAAAACACTCACT 1409
   |||
486 uSerAspIleThrSerArgPheAsnSerAlaIleGluAlaLeuAsnArg 503
   |||
1410 GCTGATATTAATACATCACTTTAATTCAGCTATTGAACACTGAACCT 1459
   |||
503 heIleGlnLysTyrAspSerValMetGlnArgLeuLeuAspAspPheSer 519
   |||
1460 TCATTCAGAAATATGATTCAGTATGACAGCAAGCTGATGACACGCT 1509
   |||
520 GlyLys 521
   |||
1510 GGTAAA 1515

seq_name: N_Geneseq_36:T38256
seq_documentation_block:
ID T38256 standard; DNA; 1462 BP.
AC T38256:
DT 28-DEC-1996 (first entry)
DE Y. pestis F1/V antigen gene fusion.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
   F1 antigen; cafi; ds.
OS Chimeric Yersinia pestis strain GB;
FH Chimeric synthetic.
FT key Location/Qualifiers
   cds 8..1450
   FT /*tag= a
   FT /product= F1/V fusion protein
   FT 452..472
   FT /*tag= b
   FT /note= "bases 452-472 is a sequence derived
   FT from PCR primers"
PD WO9628551-A1.
PN 19-SEP-1996.
PF 13-MAR-1996: G00571.
PR 13-MAR-1995: GB-005059.
PR 13-SEP-1995: GB-018946.
PR 05-DEC-1995: GB-024825.
PA (MINA ) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 96-433824/43.
DR P-PSDB; W01045.
Yersinia pestis V antigen and F1 antigen or their protective

```


DT 18-APR-1996 (first entry)
DE Partial lcrV (V antigen) gene of Y. pestis.
KW lcrV; V antigen; virulence; plague; vaccine; epitope; ss.
OS Yersinia pestis.
FH key Location/Qualifiers
FT cds 1..990
FT /tag- a
FT /note- "V antigen"
PN MO9524475-A1.
PD 14-SEP-1995.
PE 06-MAR-1995; G00481.
PR 08-MAR-1994; GB-004577.
PA (MINA) UK SEC FOR DEFENCE.
PI Leary SEC, Tildall RW, Williamson ED, Leary SE;
DR WPI: 95-328268/42.
DR P-PSDB: R79961.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PS oral or parenteral vaccines for protection against plague
CC T04222-23 are DNA sequences (lcrV) encoding all or a protective epitopic
CC part of the mature V protein of Yersinia pestis. The protein was
CC expressed as a fusion protein with maltose binding protein or
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis is
CC the highly virulent causative organism of plague in a wide range of
CC animals, including man. The V antigen (lcrV) is an unstable 37.3 kDa
CC monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen
CC is postulated to act as a virulence antigen, and transformed
CC microorganisms contg. recombinant DNA encoding a V antigen protein/
CC peptide are useful in vaccines to protect against plague.
SQ Sequence 1014 BP; 346 A; 181 C; 201 G; 286 T.

alignment_scores:
Quality: 1641.50 Length: 328
Ratio: 5.020 Gaps: 1
Percent Similarity: 99.695 Percent Identity: 99.390

alignment_block:

US-08-699-716A-2 x T04222 ..

Align seg 1/1 to: T04222 from: 1 to: 1014

```

194 G|u|p|h|e|t|i|e|a|g|a|t|y|r|g|l|u|g|l|a|s|p|r|o|c|n|h|s|p|h|e|l|e|g|u|a|s 210
    |||||
    7 GAAATTC...ATTAGAGCTACGAAACAAACCCACACATTTATTAGAGGA 53
210 p|e|u|g|l|u|s|v|a|l|a|g|u|g|l|u|e|u|t|h|r|g|l|h|s|g|y|s|e|r|v|a|l| 227
    |||||
    54 TCTGAAAAAGTTAGGTGGAGAACCACTTACGTGTCATGCTTCTTCACTT 103
227 e|u|g|l|u|e|u|a|g|l|u|e|u|v|a|l|y|s|a|p|l|y|s|a|n|i|l|a|s|p|i|l|e|s|e|r|i|e 243
    |||||
    104 TAGAAGAAATTTGCTCAGTACTGACAAAGATAAAATATAGATATTTCCATT 153
244 L|y|s|t|y|a|s|p|r|o|a|r|g|y|a|s|p|s|e|r|g|l|u|a|l|p|h|e|a|a|n|a|r|g|v|a|i|l|e|t|h 260
    |||||
    154 AAAATATGATCCCAAGAAAGATTCGAGGTTTGGCAATAGAGTATATAC 203
260 r|a|s|p|a|p|i|l|e|g|l|u|e|u|e|u|s|y|s|i|l|e|u|a|l|a|t|y|r|p|h|e|u|p|r|o|g|l|u|a 277
    |||||
    204 TGATGATATGCAATTCCTCAGAAAGAAATCTTAGCTTATTTTCAACCGAGG 253
277 s|p|h|r|i|l|e|u|e|u|s|g|l|y|l|h|s|t|y|a|s|p|a|n|g|i|n|l|e|u|g|l|a|s|n|g|i|y|l|e 293
    |||||
    254 ATGCATATCTTAAAGCGCGTATATGACAAACCACTGCAAAATGCAATC 303
294 l|y|s|a|r|g|v|a|l|y|s|g|i|u|p|h|e|u|g|i|u|s|e|r|p|r|o|a|n|h|r|g|i|n|t|r|p|e|i|u|l|e 310
    |||||
    304 AAGCGAGTAAAGAGGTTCTTGAATCATCGCGGAATACACATGGGAATT 353
310 u|a|r|g|a|l|a|p|h|e|t|i|e|a|l|a|m|e|h|a|s|p|s|e|r|e|u|h|r|a|l|a|s|p|a|r|g|i|l|e|a 327
    |||||
    354 GCGGCGTTTCATGCGAGTATGCAATTTCTTTTAAACCGCGCATGCTATCG 403

```

```

327 s|p|a|s|p|i|l|e|u|e|u|s|v|a|i|l|e|v|a|l|s|p|s|e|r|e|t|a|n|h|s|h|g|i|y|a|s|p 343
    |||||
    404 ATGATGATATTTTGAAGTGAATGATTTGATTCATCAATGAATCATCATGCTGAT 453
344 A|l|a|r|g|s|e|r|l|y|s|e|u|a|r|g|l|u|g|i|u|e|u|a|g|i|u|e|u|h|r|a|g|i|u|e|u|l|y 360
    |||||
    454 GCGCGTGAAGAGTTGCGTGAAGAAATTAGCTGAGCTTACCGCGCAATTA 503
360 s|i|l|e|r|s|e|r|v|i|l|e|g|i|n|a|g|i|u|i|e|a|s|n|l|y|s|h|i|s|e|u|s|e|r|s|e|r|g 377
    |||||
    504 GATTATTCAGTATTTCAAGCGGAATTAATTAACATCTGCTGATGCTG 553
377 l|y|h|r|i|l|e|a|n|i|l|e|h|a|s|p|l|y|s|e|r|i|l|e|a|n|e|u|e|t|a|s|p|l|y|s|a|n|l|e|u 393
    |||||
    554 GCACCATTAATATCCATGATTAATCCATTAATCTCATGAGATTA 603
394 t|y|r|g|i|y|r|t|h|r|a|p|g|i|u|g|i|l|i|e|h|e|l|y|a|l|a|s|e|r|l|a|g|i|u|r|y|l|y|s|i|l 410
    |||||
    604 TATGATTATACAGATGAAGAGATTTTAAAGCCAGCGCAGATACAA 653
410 e|l|e|u|g|i|u|s|e|r|p|r|o|c|i|n|h|r|t|h|r|i|l|e|g|i|n|v|a|l|a|s|p|g|i|s|e|r|g|i|u|t|y|s|l 427
    |||||
    654 TCTGAGAAATGGCTCAAAACCATTCAGCTGATGAGGAGGAGAGAAA 703
427 y|i|l|e|v|a|l|s|e|r|i|l|e|y|a|s|p|h|e|u|g|i|u|e|u|s|e|r|g|i|u|a|s|n|l|y|a|r|g|h|y|g|i| 443
    |||||
    704 AAATAGTCTCGATTAAGAGACTTCTTGAAAGTGAATTAAGAACCGGG 753
444 A|l|e|u|g|i|u|s|n|l|e|u|s|a|s|n|s|e|r|t|y|r|s|e|r|t|y|r|a|s|n|l|y|s|a|p|a|n|a|n|g|i 460
    |||||
    754 GCGTGGGTAATCTGAANAACCTCACTTATATATTAAGATTAATGA 803
460 u|l|e|u|s|e|r|i|s|p|h|e|a|l|a|h|r|t|h|r|c|y|s|e|r|a|s|p|l|y|s|e|r|a|r|p|r|o|l|e|u|a|s|n|a 477
    |||||
    804 ATTATTCACCTTCCACACACCTCTCGATTAATCCAGCGCGCTCAACG 853
477 s|p|l|e|u|a|l|s|e|r|g|i|n|l|y|s|t|h|r|g|i|n|l|e|u|s|e|r|a|s|p|i|l|e|h|r|s|e|r|a|r|p|h|e 493
    |||||
    854 ACTTGTTAGCCAAAACCACTCAGCTGTCTGATATTAATCATCAGTTTT 903
494 A|n|s|e|r|i|a|i|l|e|g|i|u|a|i|l|e|u|a|s|n|a|r|g|h|e|i|l|e|g|i|n|l|y|s|t|y|r|a|s|s|e|r|v|a 510
    |||||
    904 AATTCAGTATTAAGACCTGACCGTTTCATTTCAGAAATATGATTCAGT 953
510 I|m|e|r|g|i|n|a|r|g|l|e|u|a|s|p|a|s|p|h|r|s|e|r|g|i|y|l|y|s 521
    |||||
    954 GATGCAACGTCTGCTAGATGACAGTCTGTGTTAA 987

```

seq_name: N.Geneseq_36:T38242

seq_documentation_block:
ID T38242 standard; DNA: 1014 BP.
AC T38242;
DT 28-DEC-1996 (first entry)
DE Y. pestis lcrV (V antigen) gene.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
OS Yersinia pestis strain GB.
FH key Location/Qualifiers
FT cds 1..990
FT /tag- a
FT misc_feature 1..6
FT /tag- b
FT /note- "vector pMAL-p2 or pMAL-c2-derived bases"
FT mutation 1006
FT /tag- c
FT /note- "base 1006 is altered to a T to create a second in-frame stop codon"
PN MO9628551-A1.
PD 19-SEP-1996.
PE 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.

PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MTNA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR P-PSDS: W01040.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Claim 5; Page 25-28; 98PP; English.
 CC A lcrv gene sequence (T38242) codes for the Yersinia pestis V
 CC antigen (W01040), which is capable of evoking protective immune
 CC responses in animals. The gene was amplified from Y. pestis
 CC DNA by PCR using primers (T38250-51) homologous to the 5' and 3'
 CC ends of the gene. The gene was inserted into vector pMAL-P2,
 CC pMAL-c2 or pGEX-5X-2 (see also T38243) to allow prodn. of
 CC recombinant V antigen for use in vaccines against plague.
 CC Expression in gut-colonising organisms and attenuated Salmonella
 CC typhi allows prodn. of live vaccines. F1/V antigen fusions were
 CC also created (see also T38249 and T38256). The gene can itself be
 CC used in genetic vaccines.
 SQ Sequence 1014 BP; 347 A; 180 C; 201 G; 286 T;

alignment scores:
 Quality: 1641.50 Length: 328
 Ratio: 5.020 Gaps: 1
 Percent Similarity: 99.695 Percent Identity: 99.390

alignment_block:
 US-08-699-716a-2 x T38242 ..

Align seg 1/1 to: T38242 from: 1 to: 1014

194 GluphemeTlleRgAlaTyrgLugInaSprGInHsPheIleGlu 210
 |||||
 7 GAATTC..ATTGAGCCTACGACAAACCCACACATTATTATGAGG 53
 210 pleugluTyValargValgluInleuthrgLynHsglySerSerVal 227
 |||||
 54 TCTAGAAAAGTTAGGGTGAACAACCTACTGTGTCATGCTCTCAGTT 103
 227 eugluGluLeuValgluLeuValLysAspLysAsnIleAspIleSer 243
 |||||
 104 TAGAAGAAATGGTTTCAGTTAGTCAAGATAAAATATGATATTCACAT 153
 244 LysTyAspProArgLysAspSerGluValPheAlaSerArgValIle 260
 |||||
 154 AATATGATGCCCGAAGAAAGATTGGAGGTTTGGCAATAGAGTAATTC 203
 260 rAspAspIleGluLeuLeuLysLysIleLeuAlaTyrgPheLeuPro 277
 |||||
 204 TGTATGATATCGAATTCCTCAAGAAAATCCTAGCTTATTTCTACCGAG 253
 277 spThrIleLeuLysGlyGlyHisTyrgAspAsnGlnLeuGlnAsnGly 293
 |||||
 254 AAGCCATCTCTAAAGCGGTCATTATGACAACTGCAAAATGGAATC 303
 294 LysArgValLysGluPheLeuGlnSerSerProAsnThrGlnTrpGlu 310
 |||||
 304 AAGCGGTAAGAGTCTCTGATCATCGCCGAATACAAAGGGAAT 353
 310 uArgAlaPheMetAlaValMetHisPheSerLeuThrAlaAspArgIle 327
 |||||
 354 GCGGCGTTCATGCGACTAATGCATTCCTTTAACCAGCGCATGATCG 403
 327 spAspAspIleLeuLysValIleValAspSerMetAsnHisGlyAsp 343
 |||||
 404 ATGATGATATTTTGAAGTGAATGATGATCATGATCATGATGATGAT 453
 344 AlaArgSerLysLeuArgGluGluLeuAlaGluLeuThrAlaLeuLeu 360
 |||||
 454 GCCCGTAGCAAGTTGGTGAAGAAATTAAGCTGACCTTACCGCCGAAT 503

360 sIleTySerValIleGlnAlaGluIleAsnLysHisLeuSerSer 377
 |||||
 504 GATTATTCAGTATTATCAAGCGGAATTAATTAAGCATCTGCTACTAGT 553
 377 LysThrIleAsnIleHisAspLysSerSerIleAsnLeuMetAspLysAsn 393
 |||||
 554 GCACATATAATTCATGATTAATCACTTAATCTGATGATAAAATTTTA 603
 394 TyrgLysTyrgThrAspGluGluIlePheLysAlaSerAlaGluTyrg 410
 |||||
 604 TTGCGTATACGATGAGAGATTTTAAAGCCAGCCGAGATACAAAT 653
 410 eLeuGluLysMetProGlnThrThrIleGlnValAspGlySerGluLys 427
 |||||
 654 TCTCAGAAAAGTCCCTCAACACCATTCAGGTGATGGAGCGAGAAA 703
 427 yIleValSerIleLysAspPheLeuGlySerGluAsnLysArgThrGly 443
 |||||
 704 AATAGTCTCGATTAAGGACTTCTTGAGAGTGAAGATAAAGAACCGG 753
 444 AlaLeuGlyAsnLeuLysAsnSerTyrgSerTyrgAsnLysAspAsnG 460
 |||||
 754 GCGTTGGGTAATCTGAAAACCTCATCTCTTATTAATGAATATATGA 803
 460 uLeuSerHisPheAlaThrThrCysSerAspLysSerArgProLeuAsn 477
 |||||
 804 ATTACTCTACTTGGCCACCACTGCTCGGATTAAGCGCCGCTCAAG 853
 477 sPLeuValSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPhe 493
 |||||
 854 ACTTGTTAGCCCAAAACCACTGCTGATGATTAACATCACTGTTT 903
 494 AsnSerAlaIleGluAlaLeuAsnArgPheIleGlnLysTyrgSerVa 510
 |||||
 904 AATTCAGTATTAAGACCTGACCAACCTTCATTCAGAAATATGATTCAG 953
 510 lMetGlnArgLeuLeuAspAspThrSerGlyLys 521
 |||||
 954 GATGCAAGCTCTGATGATGACAGCTCTGTGTA 987

seq_name: N_Geneseq_36:T38243

seq_documentation_block:

ID T38243 standard; DNA; 1014 BP.

AC T38243:

DT 28-DEC-1996 (first entry)

DE Y. pestis lcrv (V antigen) gene.

KW Plague; vaccine; genetic immunisation; V antigen; lcrv;

KW F1 antigen; ds.

OS Yersinia pestis strain GB.

FH Key location/Qualifiers

FT cds 1..990

FT /*tag- a

FT misc.feature 1..10

FT /*tag- b

FT /*note- "vector pGEX-5X-2-derived bases"

FT mutation 16

FT /*tag- c

FT /*note- "base 16 is altered to a C from an

FT A to createan EcoRI site"

FT mutation 1006

FT /*tag- d

FT /*note- "base 1006 is altered to a T to create a

FT mutation second in-frame stop codon"

FT W09628531-A1.

PN 19-SEP-1996.

PD 13-MAR-1996; G00571.

PE 13-MAR-1995; GB-005059.

PR 15-SEP-1995; GB-018946.

PR 05-DEC-1995; GB-024825.

PA (MTNA) UK SEC FOR DEFENCE.

PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;

DR WPI: 96-433824/43.

DR P-PSDB: W01041.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 epitopic parts - useful in vaccine for protection against plague
 PS Claim 5: Page 32-35; 98pp; English.
 CC A lcrv gene sequence (T38243) codes for the Yersinia pestis V
 CC antigen (W01041), which is capable of evoking protective immune
 CC responses in animals. The gene was amplified from Y. pestis
 CC DNA by PCR using primers (T38251 and T38259) homologous to the 5'
 CC and 3' ends of the gene. The gene was inserted into vector
 CC pGEX-5x-2, pMAL-p2 or pMAL-c2 (see also T38242) to allow produ. of
 CC recombinant V antigen for use in vaccines against plague.
 CC Expression in gut-colonising organisms and attenuated Salmonella
 CC typhii allows live vaccine produ. F1/V antigen fusions were also
 CC created (see also T38249 and T38256). The gene can itself be used
 CC in genetic vaccines.
 SQ Sequence 1014 BP; 343 A; 184 C; 205 G; 282 T;

alignment_scores:
 Quality: 1641.00 Length: 325
 Ratio: 5.049 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.692

alignment_block:
 US-08-699-716A-2 x T38243

Align seg 1/1 to: T38243 from: 1 to: 1014

197 lleaarglaatyrugluinaspnproglinhispherilegluaspleuglu 213
 13 attgagaccctgacgaaacacccacacattttttgagatctcagaaaa 62
 213 svaatrgvalgluclneuthrglyhisglysersevalleuglu 230
 63 agttagggtggacacacttactgtctcattgtttcatttaacagat 112
 230 euvalglneualysaspplsasnielaspiieserileystyrasp 246
 113 tgggttcgattgacaaagatataatattagattttccatttaattgat 162
 247 proarglyaspserserigluvalphealaasnaryalilethraspapi 263
 163 cccgaaagattcgaggtttttgccaatagatatttactgatgat 212
 263 eglueneulysylsileualaityrphleuprogluaspthrile 280
 213 cgaatttcctcagaaatccatgcttattttctaccgagatcccat 262
 280 eulysglylhisstyraspasnleugluinanglylileysargval 296
 263 tttaagccgctcattatgacacacactgcacaaatggcattcaacgagta 312
 297 lysglubheugluuserseproasnthrclntrpgluueatgala 313
 313 aaagagttccttgatcatcgccgcaatacacaatgggattggcggtt 362
 313 emelalavalmethispheserleuthraliaspapiileaspaspapi 330
 363 catggcagatagcatttcttttaacccgcatcgatcatgatgata 412
 330 leleulysvalillevalasposermetasnhsisglyaspalaariser 346
 413 ttttgaagagcattgtgattcattgaattcattcattgggagcccgatgac 462
 347 lysleuargluguileualaigluethralagluueulysiletyrse 363
 463 aagttggcgaagaaattagctgacattacccgcaatttaagattatttc 512
 363 rvaillleaglaleuileasnlyshisleserseeriglythrilea 380
 513 agttattcaagccgaaattataagcattctgtcagtagtgccacacataa 562
 380 snliehisasplyserilleasnleuethasplysasnleutyrglytyr 396

|||||
 563 atatccatgataaattcattatctcattgataataaatttatatgattat 612
 397 thrasplugluillephelyasaserialglutyrlsileleuglu 413
 613 acagatagaagagatttttaagccagcagcagagacaaattctcagaa 662
 413 smetproglinthrthillegluvalaspglysergluylsilevals 430
 663 aatgctcctcaaacccattcagggatggagacagaaaaaataatgctc 712
 430 etilellyasppheleuglysergluasnlysargrthrglyalaleugly 446
 713 cgataaagacattcttgagagtagaatataaagacggcgctgggt 752
 447 asnleulysasnseryserlyasnllyaspasnanglutleuserh1 463
 763 aatctgaaaaaacatcatcttattataaagatataatattatctca 812
 463 spheialatthrthrcysserasplysersarproleuasnaspleuval 480
 813 cttggccaccacactgctcgatataagtcacagccgctcaacgacttggt 862
 480 erglnlysthrthrgluuseraspietheraserargpheasnserra 496
 863 gccaaaaaacactcagctctgataattatcattcagcttttaattcagct 912
 497 lileglualaleuasnarqphelleglnlystyrasposervalnetglnar 513
 913 attgaacacactgaacgcttattcattcagaaatattgattcagatgcacag 962
 513 gleuleaspsaspthrserylys 521
 963 tctgctagatgacacgctgtgtaaa 987

seq_name: N_Geneseq_36:T04223

seq_documentation_block:

ID T04223 standard; DNA; 1014 BP.

AC T04223;

DT 18-APR-1996 (first entry)

DE Partial lcrv (V antigen) gene of Y. pestis.

KW lcrv; V antigen; virulence; plague; vaccine; epitope; ss.

OS Yersinia pestis.

FH Key Location/Qualifiers

FT cds 1..990

FT /*tag= a

PN MO9524475-A1.

PD 14-SEP-1995.

PF 06-MAR-1995; G00481.

PR 08-MAR-1994; GB-004577.

PA (MINA) UK SEC FOR DEFENCE.

PI Leary SEC, Tiltball RM, Williamson ED, Leary SE;

DR WPI: 95-328268/42.

P-PSDB: R79962.

PT Recombinant DNA expressing Yersinia pestis V antigen - useful in

PS Claim 6: Page 15-16; 25pp; English.

CC T04222-23 are DNA sequences (lcrv) encoding all or a protective epitopic

CC expressed as a fusion protein with maltose binding protein or

CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis

CC is the highly virulent causative organism of plague in a wide range of

CC animals, including man. The V antigen (lcrv) is an unstable 37.3 kDa

CC monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen

CC is postulated to act as a virulence antigen, and transformed

CC microorganisms contg. recombinant DNA encoding a V antigen protein/
 CC peptide are useful in vaccines to protect against plague.

SQ Sequence 1014 BP; 343 A; 185 C; 204 G; 282 T;

alignment_scores:

Quality: 1638.00

Length: 325

Ratio: 5.040 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.385

Alignment block:
US-08-699-716A-2 x T04223

Align seg 1/1 to: T04223 from: 1 to: 1014

```

197 lIeArGaLaTyRGlUGlInAsnPrOgInHIsPheIlleGluAsPLeuGluLy 213
13 ATTCGAGCCTACGAAACCAACCAACATTTTATGAGATCTAGAGAAA 62
213 sValArVaIcIuGlIneUthrGlyHISGlySerSeRValleuGluGlu 230
63 AGTAGGGTGGAGACACTTACTGGTCATGCTTCTTCACTTTAGAGAAAT 112
230 euValGInleuVallyAspLySAsnIleAspIleSerIleLySTyrSP 246
113 TGCTTCAGTTAGTCAGAAATATAATAGATATTCCATTAATATGAT 162
247 ProArGlyAspSerGluValAlPheAlaAsnArGValIleThrAspAspI 263
163 CCCAGAAAAGATTCGAGGTTTTCGCAATAGAGTAATTAAGATGATAT 212
263 eGluLeuLeuLySlyIleleuAlaTyRPhelLeuProGluAspThrIle 280
213 CGAATTCCTCAAGAAATCCTAGCTATTCTTCAAGATGATGATGATC 262
280 euLySGlyGlyHISTyRAspAsnGluGluAsnGlyIleLySArGVal 296
263 TTTAAAGGGGCTATTATGACAAACCACTGCAAAATGGATCAAGCGACTA 312
297 LySGluPhelLeuGluSerSerProAsnThrGlnTrpGluLeuArGAlaPh 313
313 AAAAGATTCCTGATATCATCCGCAATACACATGCGAATTCGCGGCTT 362
313 eMeAlaValMetHISpHeSerIleUthrAlaAspArGlyIleAspAspI 330
363 CATGCGCATGATGATTCCTTAAACCGCGATCGATCATGATGATGAT 412
330 lIeUlyValIleValAspSerMetAsnHISGlyAspAlaArGer 346
413 TTTGAAAGTATGATTGATTCATATGATCATGATGATGATGATGATG 462
347 LySeuArGglUGluLeuAlaGluLeuThrAlaGluLeuLySlyIleTyrSe 363
463 AAGTTCGCTGAAGATTAAGCTCAGCTTACCCCGAATTAAGATTTATTC 512
363 lValIleGluAlaGluIleAsnLySHISLeuSerSerSerGlyThrIleA 380
513 AGTATTCACACCGAAATTAATAGCATCTGCTAGTAGTGACACCATAA 562
380 snIleHISAspLySerIleAsnLeuMetAspLySAsnLeuTyGlyYr 396
563 ATATCCAGATTAATCCATTAATCATGAGTAAATAATTTATATGATGAT 612
397 ThrAspGluGluIlePheLySAlaSerAlaGluTyRlySlyIleGluGly 413
613 ACAGATGAAGATTTTAAAGCCAGCCAGAGTACAAATATTCGAGAA 662
413 sMeTProGInThrThrIleGluValAspGlySerGluLySlyIleValS 430
663 AATGCTCAACCAACCATTCAGTGGATGGAGGAGAGAAAATAATAGTCT 712
430 erIleLySAspPheLeuGlySerGluAsnLySArGThGlyAlaLeuGly 446
713 CGATTAAGAGACTTCTTGGAGTGAAGTAAAGAACCGGGCGCTGGGT 762
447 AsnLeuLySAsnSerTyRSerTyRAsnLySAspAsnAsnGluLeuSerH 463
763 AATCTGAAAACTCATCTCTTATATATAAGTAAATATATGATTTATCTCA 812
463 sPheAlaThrThyRysSerAspLySArGProLeuAsnAspLeuValS 480

```

```

|||||
813 CTTGCCACACCGCTCGGTAGATGTCAGCCGCTCAACGACTTGTTA 862
480 erGInLySThrThrGluLeuSerAspIleThrSerArGpHeAsnSerAla 496
863 GCCAAAACACTCAGCTGTGATTAATTAACATCAGCTTTAATTCAGCT 912
497 lIeGluAlaLeuAsnArGpHeIleGInLySTyrAspSerValMetGlnAr 513
913 ATTAAGACACTGAACCGTTTCATTGAGAAATATGATCATGATGACAGC 962
513 gLeuLeuAspAspThrSerGlyLyS 521
963 TCTCTAGATGACACGCTGTGTA 987

```

seq_name: N_Geneseq_36:Q92819

```

seq_documentation_block:
ID Q92819 standard; DNA; 542 BP.
AC Q92819;
DR 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen in plasmid pPoreFb.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typh;
KM bubonic plague; pneumonic plague; ds.
OS Yersinia pestis.
FH key
FT Location/Qualifiers
FT 2..7
FT /*tag= a
FT /note= "first protein encoded by pPoreFb"
FT 1..6
FT /*tag= b
FT misc_feature
FT 536..541
FT /*tag= c
FT /note= "cafi open reading frame downstream seq."
FT 21..530
FT /*tag= d
FT /*tag= "cafi fusion protein"

```

```

FT WO9518231-A1.
FT 06-JUL-1995.
FT 23-DEC-1994; G02818.
FT 24-DEC-1993; GB-026425.
FT (MINA ) UK SEC FOR DEFENCE.
FT PA Howells A, Leary SEC, Oyston PCF, Tildall RW, Williamson ED;
FT WPI: 95-246396/32.
FT DR P-PSDB: R76528.
FT DNA constructs capable of transforming microorganisms - which can be
FT used as live or attenuated vaccines which induce an immune response.
FT PT against Yersinia pestis, at mucosal surfaces.
FT PS Disclosure; Page 19-20; 27pp; English.
FT CC The sequence represents the plasmid pPoreFb including the entire
FT Y. pestis cafi (F1) antigen gene having a 5' tail including a SacI
FT restriction site, and up to 1474G downstream of the cafi ORF.
FT CC The DNA construct can be used to transform human or animal gut
FT CC colonizing microorganisms, specifically attenuated Salmonella
FT CC typhimurium or Salmonella typh. The transformed microorganisms
FT CC can be used as live/attenuated vaccines which induce immune
FT CC responses at mucosal surfaces. The vaccines provide protection
FT CC against infection with Y. pestis, and are parenterally and orally
FT CC active vaccines offering protection against bubonic and pneumonic
FT CC plague.
FT SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

```

alignment_scores:
Quality: 852.00 Length: 171
Ratio: 4.982 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.415

alignment_block:
US-08-699-716A-2 x Q92819

Align seg 1/1 to: Q92819 from: 1 to: 542

```

23 HisMetLysLysIleSerSerValIleAlaIleAlaLeuPheGlyThrI1 39
18 TATATGAAAAAATCATGTCCTTATCGCCATTCATTAATTTGGAACTAT 67
39 eAlaThrAlaAsnAlaAlaAspLeuThrAlaSerThrThrAlaThrAlaT 56
68 TGCACACTGCTAATGCGCAGATTTACTGCAAGCACCACCTGCAAGCGCA 117
56 hLeuValGluProAlaArgIleThrLeuThrTyrLysGluGlyAlaPro 72
118 CTCTGTTGTAACCGAGCCGATCATCTTACATATATAGAAAGCGCTTCA 167
73 IleThrIleMetAspAsnGlyAsnIleAspThrGluLeuValGlyTh 89
168 ATTACAAATTATGACATGAAACATGATACAGATTCATCTGTTGATC 217
89 rLeuThrLeuGlyLysLysThrGlyThrThrSerThrSerValAsp 106
218 GCTTACTCTGGCGGCTATTAACAGAACACCTAGCACATCTGTTAACT 267
106 hEThrAspAlaAlaGlyAspProMetTyrLeuThrPheThrSerGlnAsp 122
268 TTACAGATGCGCGGCTGATCCATCTACTTAACATTTACTTCTCAGAT 317
123 GlyAsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArgAs 139
318 GGAATAAACACCAATTCACATAAAAGATGGCAAGATTCTAGAGA 367
139 pPheAspIleSerProLysValAsnGlyLysAsnLeuValGlyAspAsp 156
368 TTTTGTATATCTCTCTTAAGTAAACGCTAGAGAACTTGtGGGGATGAGC 417
156 aLValLeuAlaThrGlySerGlnAspPhePheValArgSerIleGlySer 172
418 TCGTCTTGCTAGCGGACAGCATTTCTTGTTCGCTCAATGGTTC 467
173 LysGlyGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 189
468 AAAGCGGTAAACTGACAGAGTAATACACTGATGCTGTAACCGTAAAC 517
189 rValSerAsnGln 193
518 CGTATCTAACCA 530

seq_name: N_Geneseq_36:V41594

seq_documentation_block:
ID V41594 standard; DNA; 544 BP.
AC V41594;
DT 26-OCt-1998 (first entry)
DE Nucleotide sequence of F1 antigen nyf1(a)sec544.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH Key location/Qualifiers
FT CDS 17..529
FT /*tag= a
FT /product= "F1 antigen"
PN WO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EV, Osorio JE, Thomas RE;
DR WPI: 98-33331/29.
PT Protection of animals against plague - using nucleic acid encoding
PS antigen from Yersinia, Pasteurella and Francisella spp.
CC Claim 8; Page 51; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.

```

```

SO Sequence 544 BP; 164 A; 120 C; 114 G; 146 T;

alignment_scores:
Quality: 852.00 Length: 171
Ratio: 4.982 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.415

alignment_block:
us-08-699-716a-2 x V41594

Align seg 1/1 to: V41594 from: 1 to: 544

23 HisMetLysLysIleSerSerValIleAlaIleAlaLeuPheGlyThrI1 39
18 TATATGAAAAAATCATGTCCTTATCGCCATTCATTAATTTGGAACTAT 67
14 TATATGAAAAAATCATGTCCTTATCGCCATTCATTAATTTGGAACTAT 63
39 eAlaThrAlaAsnAlaAlaAspLeuThrAlaSerThrThrAlaThrAlaT 56
68 TGCACACTGCTAATGCGCAGATTTACTGCAAGCACCACCTGCAAGCGCA 113
56 hLeuValGluProAlaArgIleThrLeuThrTyrLysGluGlyAlaPro 72
114 CTCTGTTGTAACCGAGCCGATCATCTTACTATATAGAAAGCGCTTCA 163
73 IleThrIleMetAspAsnGlyAsnIleAspThrGluLeuValGlyTh 89
164 ATTACAAATTATGACATGAAACATGATACAGATTCATCTGTTGATC 213
89 rLeuThrLeuGlyLysLysThrGlyThrThrSerThrSerValAsp 106
214 GCTTACTCTGGCGGCTATTAACAGAACACCTAGCACATCTGTTAACT 263
106 hEThrAspAlaAlaGlyAspProMetTyrLeuThrPheThrSerGlnAsp 122
264 TTACAGATGCGCGGCTGATCCATCTACTTAACATTTACTTCTCAGAT 313
123 GlyAsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArgAs 139
314 GGAATAAACACCAATTCACATAAAAGATGGCAAGATTCTAGAGA 363
139 pPheAspIleSerProLysValAsnGlyLysAsnLeuValGlyAspAsp 156
364 TTTTGTATATCTCTCTTAAGTAAACGCTAGAGAACTTGtGGGGATGAGC 413
156 aLValLeuAlaThrGlySerGlnAspPhePheValArgSerIleGlySer 172
414 TCGTCTTGCTAGCGGACAGCATTTCTTGTTCGCTCAATGGTTC 463
173 LysGlyGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 189
464 AAAGCGGTAAACTGACAGAGTAATACACTGATGCTGTAACCGTAAAC 513
189 rValSerAsnGln 193
514 CGTATCTAACCA 526

seq_name: N_Geneseq_36:V41596

seq_documentation_block:
ID V41596 standard; DNA; 544 BP.
AC V41596;
DT 26-OCt-1998 (first entry)
DE Nucleotide sequence of F1 antigen nyf1sec510.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH Key location/Qualifiers
FT CDS 17..532
FT /*tag= a
FT /product= "F1 antigen"
PN WO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.

```

PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haenes EJ, Osorio JE, Thomas RE;
 DR WPI; 98-333331/29.
 DR P-PSDB; W59783.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Pages 33-34; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 344 BP; 166 A; 118 C; 112 G; 148 T;

alignment_scores:
 Quality: 852.00 Length: 171
 Ratio: 4.982 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.415

alignment_block:
 US-08-699-716a-2 x V41596

Align seg 1/1 to: V41596 from: 1 to: 544

```

23 Hismetlyslsileserservalillealialealeupheglythrll 39
14 TATATGAAAAAATCAGTTCCTTCGATCCGATTCGATTCATTTGGAACTAT 63
39 ealathralaasnaalalaaspleuthralaaserthrthralathralat 56
64 TGCACATGCAATAGCGGACAGATTACTGCACACACACGCAAGCGCA 113
56 hrlenuvalguiproalargillethleuthrlyrlysglugialapro 72
114 CTCTGTGTGACACGCCGCGATCCTTACATATTAAGAGGCGCTCCA 163
73 lilethrilemetaspasnnglyasniileaspthrGleuenuvalglyth 89
164 ATTACATTTTGACACATGGAACATCGATACACAAATTACTTGTGGTAC 213
89 rleuthrleuglyglyrlysthrlythrlythrlythrlythrlythrly 106
214 GCTTACTCTGGCGGCGATTAAGAGACACACACATCTGTTACT 263
106 hehraspalaalaglyaspProwmetrlyrleuthrphetrserglasp 122
264 TTACAGATGCGGCGGATCCATGATCTTACATTTACTTCTCAGAGAT 313
123 G1YAsnaahisglnphetrthrlythrlythrlythrlythrlythrly 139
314 GGAATATACCAACCAATTCATCAAAAGTATGCGAAGATCTTAAGA 363
139 pheaspilseerProlyvalasnnglyluasnleuvalglyaspaspv 156
364 TTTTGAATATCTCTCCTAAGGTAAAGGTGAACCTTGTGGGGATACG 413
156 alvalleualathnglyserglaspPhevalargserilleglyser 172
414 TCCTCTGCGTACGCGGACGACAGATTTCTTGTCTCCTCAATGTGTCC 463
173 Lysglyglylystleuualaglylystlythraspalaalathrvalth 189
464 AAAGGCGGTAACTGCAGCAGTAATATACATGATCTGAACCGTAAC 513
189 rValserasnngln 193
514 CGTATCTAACCA 526
seq_name: N_Geneseq_36:T38248
seq_documentation_block:

```

ID T38248 standard; DNA; 547 BP.
 AC T38248;
 DE 28-DEC-1996 (first entry)
 DR Y. pestis F1 antigen caf1 gene (including signal sequence).
 KW Plague; vaccine; genetic immunisation; F1 antigen; caf1;
 V antigen; ds.
 OS Yersinia pestis strain GB.
 FH Key Location/Qualifiers
 FT signal_peptide 29..91
 FT mat_peptide 92..538
 FT /*tag= a
 FT /*tag= b
 FM W0628551-A1.
 PD 19-SEP-1996.
 PR 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MVA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
 DR WPI; 96-433824/43.
 DR P-PSDB; W01043.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PI epitopic parts - useful in vaccine for protection against plague
 PS Claim 41; Page 61-62; 98pp; English.
 CC A DNA sequence (T38248) comprises the caf1 gene, including the
 CC signal sequence, coding for the F1 antigen (W01043) of Yersinia
 CC pestis. It was obt. by PCR amplification (see also T38257-58)
 CC of Y. pestis DNA. The PCR product was cloned into plasmid pBKCMV
 CC and the resulting plasmid (pF1AB) was used to transform E. coli Nova
 CC Blue. Purified plasmid, when administered by i.m. injection,
 CC induced an immunoglobulin response to F1 in Balb/c mice. Live
 CC vaccines comprising gut colonising organisms transformed with the
 CC caf1 gene (see also T38244) can be used to protect a host animal
 CC against plague.
 SQ Sequence 347 BP; 165 A; 120 C; 115 G; 147 T;

alignment_scores:
 Quality: 852.00 Length: 171
 Ratio: 4.982 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.415

alignment_block:
 US-08-699-716a-2 x T38248

Align seg 1/1 to: T38248 from: 1 to: 547

```

23 Hismetlyslsileserservalillealialealeupheglythrll 39
26 TATATGAAAAAATCAGTTCCTTCGATCCGATTCGATTCATTTGGAACTAT 75
39 ealathralaasnaalalaaspleuthralaaserthrthralathralat 56
76 TGCACATGCAATAGCGGACAGATTACTGCACACACACGCAAGCGCA 125
56 hrlenuvalguiproalargillethleuthrlyrlysglugialapro 72
126 CTCTGTGTGACACGCCGCGATCCTTACATATTAAGAGGCGCTCCA 175
123 lilethrilemetaspasnnglyasniileaspthrGleuenuvalglyth 89
176 ATTACATTTTGACACATGGAACATCGATACAGAAATTACTTGTGGTAC 225
89 rleuthrleuglyglyrlysthrlythrlythrlythrlythrlythrly 106
226 GCTTACTCTGGCGGCTATTAAGAGACAGACACATCTGTTACT 275
106 hehraspalaalaglyaspProwmetrlyrleuthrphetrserglasp 122
276 TTACAGATGCGGCGGATCCATGATCTTACATTTACTTCTCAGAGAT 325
123 G1YAsnaahisglnphetrthrlythrlythrlythrlythrlythrly 139

```

```

|||||
326 GGAATATACCAACCAATTCACATACAAAGTGAATGGCAGAGATTCAGAGA 375
139 pPhaspiIleSerProlySValasnGlyGluasnLeuValGlyaspav 156
376 TTTTGATATCTCTCTCTAAGGTAAACGGTGAAGACCTTGTGGGGGTGACG 425
156 aValleuAlaThrGlySerGlnaspPheValArgSerIleGlySer 172
426 TCGCTTGCTGACGGGACGACGAGATTTCTTGTGCTCATGTGCTCC 475
173 LysGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 189
476 AAAGCGGTAAACTTGGCAGGATAAATACACTGATGCTGTATACCGTAAC 525
189 rValSerAsnGln 193
526 CGTATCTAACCAA 538

```

seq_name: N_Geneseq_36.V41595

seq_documentation_block:

```

ID V41595 standard; DNA; 510 BP.
AC V41595;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of FI antigen nYpI(b)sec544.
KW FI antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
PN MO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
PT Protection of animals against plague - using nucleic acid encoding
PS antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 53; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis FI antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 510 BP; 153 A; 114 C; 106 G; 137 T;

```

alignment_scores:

```

Quality: 850.00 Length: 170
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-08-699-716A-2 x V41595 ..

Align seg 1/1 to: V41595 from: 1 to: 510

```

24 MetLysLysIleSerSerValIleAlaIleAlaLeuPheGlyThrIleAl 40
|||||
1 ATGAAAAAAACATCGATTCGTCATTCATTCATTCATTCATTCATTCATTC 50
40 aThrAlaAsnAlaAlaAspLeuThrAlaSerThrThrAlaThrAlaThrL 57
|||||
51 AACGCTAAATCGCGAGATTAACTGCAAGCACACCTGCACAGCGCAACTC 100
57 euValGluProAlaArgIleThrLeuThrTyrLysGluGlyAlaProIle 73
|||||
101 TTGTGTAACCAAGCCGATCATCTTACATATAAGAAAGCGCTCCATT 150
74 ThrIleLeuAspAsnGlyAsnIleAspThrGluLeuLeuValGlyThrLe 90
|||||
151 ACAATATATGACATGGAACATCATGATACAGAAATATTACTTGTGGAGCT 200
90 uThrLeuGlyGlyTyrIleThrGlyThrThrSerThrSerValAsnPheT 107

```

```

|||||
201 TACTCTTGCGGCTATTAACAGAACACCTACGACATCTTACTTTA 250
107 hrAspAlaAlaGlyAspPheMetTyrLeuThrPheThrSerGlnaspGly 123
251 CAGATGCCGGGGTGATGCCATGTACTTAACATTACTTCTCAGAGAGA 300
124 AsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArgAspPh 140
301 AATAACCAACCAATTCATACAAAGTATGTCAGAGATTCCTAAGATTT 350
140 eAspIleSerProlySValasnGlyGluasnLeuValGlyaspav 157
351 TGATATCTCTCTTAAGGTAAACGGTGAAGACCTTGTGGGGATGACGTCG 400
157 aLeuAlaThrGlySerGlnaspPhePheValArgSerIleGlySerLys 173
401 TCTTGCTGACGGGACGACGAGATTTCTTGTGCTCATGTGCTCCAAA 450
174 GlyGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 190
451 GCGGTAAACTTGCAGCAGGATAAATACACTGATGCTAACCCTAACCGT 500
190 lSerAsnGln 193
501 ATCTAACCAA 510

```

seq_name: N_Geneseq_36.V41609

seq_documentation_block:

```

ID V41609 standard; DNA; 447 BP.
AC V41609;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of FI antigen nYpImat447.
KW FI antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH Key 1.447
FT Location/Qualifiers
FT CDS
PN MO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR P-PSDB; W59788.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 64; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis FI antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 447 BP; 133 A; 102 C; 95 G; 117 T;

```

alignment_scores:

```

Quality: 755.00 Length: 149
Ratio: 5.067 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-08-699-716A-2 x V41609 ..

Align seg 1/1 to: V41609 from: 1 to: 447

```

45 AlaAspLeuThrAlaSerThrThrAlaThrLeuValGluProAl 61
|||||
1 GCAGATTTAACTGCAAGCACACCTGCACGCAACTCTTGTGTAACAGC 50

```



```

61 aargllethleuthrtyrlysgluliyalaProilethrillemetaspA 78
|||||
51 CGGCATCACTCTTACATATAAGAGAGCGCTCCAAATTAATATAGACA 100
|||||
78 snglyAsnIleasphthrguleuleuValgIYthrlleuthrleuglygly 94
|||||
101 AVGGAAACATCATACAGAAATTAATTGTTGTAACGCTTCTTGCGCGC 150
|||||
95 TyrlysthrGlyThrThrSerThrSerValaenphethrAspAlaAlaGl 111
|||||
151 TATATAAACGGAACCACTAGCACATCTGTAACTTACAGATGCCGCGG 200
|||||
111 YASpPromeTYrleuthrPheThrSerGlnaspGlyAsnAsnHISGlnP 128
|||||
201 TGATCCCACTACTTAACATTTACTTCTCAGATGGAATAACCAACCAAT 250
|||||
128 heThrThrlYsValIleGlyLysAspSerArgAspPheAspIleSerPro 144
|||||
251 TCACATCAAAAGTAGATTGGCAAGATTCTAGAGATTGTGATATCTCTCT 300
|||||
145 lysValAsnGlyGlnAsnleuValGlyAspAspValValleuAlaThrGcl 161
|||||
301 AAGGTAAAGGGTAGAACCTTGTGGGGATGACGTCTGTGGCTACGGG 350
|||||
161 yserGlnAspPhePheValArgSerIleGlySerIlysglyLysLeuA 178
|||||
351 CAGCCGAGATTCTTGTTCGCCCAATGGTTCCAAAGCGGTAACCTTG 400
|||||
178 lAlaGlyLysTYrThrAspAlaValThrValThrValSerAsnGln 193
|||||
401 CAGCAGATAATACACTGATGCTGTACCGTAACGTAATCTAACCAA 447
|||||

```

seq_name: N_Geneseq_36:V41601

```

seq_documentation_block:
ID V41601 standard; DNA: 450 BP.
AC V41601:
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nYF1mat450.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
PN MO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 61; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 450 BP; 134 A; 102 C; 96 G; 118 T;

```

alignment_scores:
 Quality: 755.00 Length: 149
 Ratio: 5.067 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-699-716A-2 x V41601 ..
 Align seg 1/1 to: V41601 from: 1 to: 450

```

45 AlaAspleuthrAlaSerThrThrAlaThrAlaThrlleuValGluProAl 61
|||||
4 GCAGATTTAATCAAGACCACTGCAAGCGCAACTCTTGTGAAACGAGC 53
|||||

```

```

61 aargllethleuthrtyrlysgluliyalaProilethrillemetaspA 78
|||||
54 CGGCATCACTCTTACATATAAGAGAGCGCTCCAAATTAATATAGACA 103
|||||
78 snglyAsnIleasphthrguleuleuValgIYthrlleuthrleuglygly 94
|||||
104 AVGGAAACATCATACAGAAATTAATTGTTGTAACGCTTCTTGCGCGC 153
|||||
95 TyrlysthrGlyThrThrSerThrSerValaenphethrAspAlaAlaGl 111
|||||
154 TATATAAACGGAACCACTAGCACATCTGTAACTTACAGATGCCGCGG 203
|||||
111 YASpPromeTYrleuthrPheThrSerGlnaspGlyAsnAsnHISGlnP 128
|||||
204 TGATCCCACTACTTAACATTTACTTCTCAGATGGAATAACCAACCAAT 253
|||||
128 heThrThrlYsValIleGlyLysAspSerArgAspPheAspIleSerPro 144
|||||
254 TCACATCAAAAGTAGATTGGCAAGATTCTAGAGATTGTGATATCTCTCT 303
|||||
145 lysValAsnGlyGlnAsnleuValGlyAspAspValValleuAlaThrGcl 161
|||||
304 AAGGTAAAGGGTAGAACCTTGTGGGGATGACGTCTGTGGCTACGGG 353
|||||
161 yserGlnAspPhePheValArgSerIleGlySerIlysglyLysLeuA 178
|||||
354 CAGCCGAGATTCTTGTTCGCCCAATGGTTCCAAAGCGGTAACCTTG 403
|||||
178 lAlaGlyLysTYrThrAspAlaValThrValThrValSerAsnGln 193
|||||
404 CAGCAGATAATACACTGATGCTGTACCGTAACGTAATCTAACCAA 450
|||||

```

seq_name: N_Geneseq_36:V41600

```

seq_documentation_block:
ID V41600 standard; DNA: 474 BP.
AC V41600:
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nYF1mat474.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
PN MO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Pages 59-60; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 474 BP; 143 A; 106 C; 99 G; 126 T;

```

alignment_scores:
 Quality: 755.00 Length: 149
 Ratio: 5.067 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-699-716A-2 x V41600 ..
 Align seg 1/1 to: V41600 from: 1 to: 474

```

45 AlaSpleuThrAlaSerThrThrAlaThrAlaThrLeuValGluProAl 61
|||||
10 GCACATTAACTGCAGACACCACTGCACCACTGTTGTAACCAAGC 59
61 ArgIleThrLeuThrTyrLysGluGlyAlaProIleThrIleLeuAspA 78
|||||
60 CCGCATCAGCTTACATATAAGAGAGCGCCCATTAATTAATTTGGACA 109
78 snGlyAsnIleAspThrGluLeuLeuValGlyThrLeuThrLeuGlyGly 94
|||||
110 ATGGAACACATGATACAGAAATTAATTTGTTAGCTTACTCTTGGCGGC 159
95 TyrLysThrGlyThrThrSerThrSerValAsnThrAspAlaAlaGly 111
|||||
160 TATTAACAGACAGACACTGACATCTGTTACTTACAGATGCGCGGG 209
111 YAspProMetTyrLeuThrPheThrSerGlnAspGlyAsnAsnHsGlnP 128
|||||
210 TGATCCATGACTTAACATTAATTAATTTGTTAGCTTACTCTTGGCGGC 259
128 heThrThrLysValIleGlyLysAspSerArgAspPheAspIleSerPro 144
|||||
260 TCACTACAAAAGTATGCGAAGATTTAGATTTGATATCTCTCCCT 309
145 LysValAsnGlyGluAsnLeuValGlyAspAspValValLeuAlaThrGly 161
|||||
310 AAGGTAAACGGTGAACCTGTTGGGGATGACCTCGTTGGCTACGGG 359
161 YSerGlnAspPhePheValArgSerIleGlySerLysGlyGlyLysLeuA 178
|||||
360 CAGCAGAGATTCTTGTTCCTCAATTTGCTTCAAAAGCGGTAAACTTG 409
178 AlaGlyLysTyrThrAspAlaValThrValThrValSerAsnGln 193
|||||
410 CAGCAGCTAAATACATGATGCTTAACCGTAACCGTATTAACCAA 456
seq_name: N_Geneseq_36:Q92817

```

```

seq_documentation_block:
ID 092817 standard; DNA; 541 BP.
AC 092817;
DE 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen in plasmid pFCA12A.
KM Vaccine; antigen; Salmonella typhimurium; Salmonella typh.;
KM bubonic plague; pneumonic plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT cds 2..454
FT FT /*tag= a
FT FT misc_feature 1..6
FT FT /*tag= b
FT FT misc_feature 536..541
FT FT /*tag= c
FT FT /*note= "vector pFCA12A bases"
PK MO9518231-A1.
PD 06-JUL-1995.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA) UK SEC FOR DEFENCE.
PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
PI WPI: 95-246396/32.
DR P-ESDB: R76526.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.
PS Claim 7; Page 15-16; 27pp; English.
CC The sequence represents the plasmid pFCA12A construct showing the
CC fusion of the first few bases of beta-galactosidase in the vector
CC with the Y. pestis cafi (F1) antigen minus its signal sequence and
CC having a 5' tail including a SacI restriction site, and up to the
CC cafi AACCC-3' end with some vector bases. The DNA construct can be

```

CC used to transform human or animal gut colonizing microorganisms,
 CC specifically attenuated Salmonella typhimurium or Salmonella typh.
 CC The transformed microorganisms can be used as live/attenuated
 CC vaccines which induce immune responses at mucosal surfaces. The
 CC vaccines provide protection against infection with Y. pestis, and
 CC are parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 541 bp; 163 A; 120 C; 111 G; 147 T;

alignment_scores:
 Ratio: 755.00 Length: 151
 Percent Similarity: 100.000 Gaps: 0
 Percent Identity: 98.013

alignment_block:
 US-08-699-716a-2 x Q92817

Align seg 1/1 to: Q92817 from: 1 to: 541

```

43 AsnAlaAlaSpleuThrAlaSerThrThrAlaThrAlaThrLeuValGly 59
|||||
2 AGCTCGGACAGATTAACTGCAAGACACCACTGCACCACTGTTGTA 51
59 UPProAlaArgIleThrLeuThrTyrLysGluGlyAlaProIleThrIle 76
|||||
52 ACCGACCCGATCATCTATTAATTAAGAGAGCGCCCATTAACAATTA 101
76 eAspAsnGlyAsnIleAspThrGluLeuLeuValGlyThrLeuThrLeu 92
|||||
102 TGGCAATGGAACATGATACAGAAATTAATTTGTTAGCTTACTCTT 151
93 GlyLysTyrLysThrGlyThrThrSerThrSerValAsnThrAspAla 109
|||||
152 GCGGCTATTAACAGACAGACACTGACATCTGTTAACTTACAGATGC 201
109 AlaGlyAspProMetTyrLeuThrPheThrSerGlnAspGlyAsnAsn 126
|||||
202 CCGCGGTGATCCCATGATTAATTAATTTGCTTCAAAAGCGGTAACT 251
126 LysGlnPheThrThrLysValIleGlyLysAspSerArgAspPheAspIle 142
|||||
252 ACCAATTCACCTACAAAAGTATGCGAAGATTTCTAGAGATTTGATATC 301
143 SerProLysValAsnGlyGluAsnLeuValGlyAspAspValValLeuAl 159
|||||
302 TCTCCTAAGGTAAACGGTGAACCTTGTGGGGATGACGCTCGTGGC 351
159 aThrGlySerGlnAspPhePheValArgSerIleGlySerLysGlyGly 176
|||||
352 TACGGCAGCGCAGGATTCTTGTTCCTCAATTTGTTCCAAAGCGGTA 401
176 YsLeuAlaAlaGlyLysTyrThrAspAlaValThrValThrValSerAsn 193
|||||
402 AACTTGCACAGATTAATTAACATGATGCTTAACCGTATTAAC 451
193 Gln 193
|||
452 CAA 454

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Percent Similarity: 99.004 Percent Identity: 98.606

alignment_block:

US-08-699-716a-2 x US-08-913-477-22

Align seg 1/1 to: US-08-913-477-22 from: 1 to: 1530

```

23 HisMetLysLysLysSerSerValIleAlaIleAlaLeuPheGlyThrI1 39
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
10 TATATGAAAAAATCAGTCCGTTATCCCATTCGATTTATTTGGAAACAT 59
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
39 eAlaThrAlaAsnAlaAlaAspLeuThrAlaSerThrAlaThrAlaT 56
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
60 TGCACAGCTGCTATCGCGAGATTTAAGTCAAGACCACTGCAACGGCAA 109
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
56 hLeuValGluProAlaArgIleThrLeuThrIleGlyGluAlaPro 72
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
110 CTCTGTTGAAACAGCCGCATCTTACATATTAAGAAAGCGGCTCCA 159
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
73 IleThrIleMetAspAsnGlyAsnIleAspThrGluLeuValGlyTh 89
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
160 ATTACAAATTATGACAAATGAAACATCATACAGATTTCTTGTGTAC 209
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
89 rLeuThrLeuGlyLysIleThrGlyThrGlyThrSerThrSerVala 106
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
210 GCTTACTCTTGCGGCTATTAACAGAACCACTAGCACATCTGTTACT 259
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
106 hethrAspAlaAlaGlyAspProMetThrLeuThrPheThrSerGln 122
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
260 TTACAGATGCGCGGCTATCCCATGTTACTTACATTTCTTCAGAGAT 309
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
123 GlyAsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArg 139
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
310 GGAATATACCCACCAATCTACTACAAAGATGTGGCAAGATCTTAGA 359
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
139 pPheAspIleSerProLysValaAsnGlyLysLeuValGlyAspAsp 156
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
360 TTTTGATATCTCTCTAAGGTAAACGGTGAACCTTGGGGGATGACG 409
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
156 aValIleAlaIleThrGlySerGlnAspPhePheValArgSerIle 172
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
410 TCGTCTGGTACGCGGACCGAGATTTCTTGTGCTCAATTTGTTCC 459
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
173 LysGlyGlyLysLeuAlaAlaGlyLysIleThrAspAlaValIleTh 189
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
460 AAAGCGGTAACTGTCAGAGGTAAATACATGCTGTACCGGTAC 509
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
189 rValSerAsnGln.....GluPheMetIleArgAlaThrGluGln 203
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
510 CGTATCTTACCAAGATCCATCGAAGGTGCTATTAGACCTTAGACAA 559
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
203 snProGlnHisPheIleGluAspLeuGluLysValArgValGluGln 219
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
560 ACCCAACAACATTTATTGAGATCTAGAAAAGTAGGGTGGAACAATT 609
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
220 ThrGlyHisGlySerSerValLeuGluGluLeuValGlnLeuValAs 236
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
610 ACTGCTCATGTTCTTCAGTTTAGAAGAAATGTTGCTCAGTTAGTCA 659
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
236 pLysAsnIleAspIleSerIleLysThrAspProArgLysAspSerG 253
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
660 TAAAAATATGATATTTCCATTAATATGATCCCAAAAAAATTCGGAG 709
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
253 aPheAlaAsnArgValIleThrAspAspIleGluLeuLysLysIle 269
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
710 TTTTGGCAATAGAGTATTAAGTATGATATCGAATTCCTCAAGAAATC 759
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
270 LeuAlaIlePheLeuProGluAspThrIleLeuLysGlyHisIleTh 286
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
760 CTACTATTATTTCTACCCGAGAGTCCATCTTAAAGCGGTCTATTAGA 809
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
286 pAsnGluLeuGlnAsnGlyIleLysArgValLysGluPheLeuGlnS 303
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

```

810 CAACCACTGCAAAATGCAATCAAGCAGTAAAGACTTCCTTCATCAT 859
303 erProAsnThrGlnTrpGluLeuArgAlaPheMetAlaValMetHisP 319
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
860 CGCCGAATACACAAATGGGAATTCGGGCGTTCATGCGAGATTAATCAT 909
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
320 SerLeuThrAlaAspArgIleAspAspAspIleLeuLysValIleVal 336
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
910 TCTTTAACCCCGCGATCGTATCGATGATGATATTTTGAAGATGATTG 959
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
336 pSerMetAsnHisIleGlyAspAlaArgSerLysLeuArgGluLeuA 353
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
960 TTCAATGATCATCATGCTATGATGCCCTACCACTTCGTCGAACATTA 1009
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
353 lAgIleuThrAlaGluLeuLysIleThrSerValIleGlnAlaGluIle 369
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1010 CTGAGCTTACCGCCGAATTAAGATTATTATTCAGTTATTCAAGCCGA 1059
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
370 AsnLysHisLeuSerSerSerGlyThrIleAsnIleHisAspLysSe 386
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1060 AATAGCATCTGTCTAGTAGGCGACCATTAATATCATATTAATTCAT 1109
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
386 eAsnLeuMetAspLysAsnLeuIleThrGlyThrAspGluGluIleP 403
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1110 TAACTCATGATATAAAATTTATATGTTATACAGATGAAGAGATTTT 1159
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
403 yAlaSerAlaGluTrpLysIleLeuGluLysMetProGlnThrIle 419
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1160 AAGCAGCGGAGAGTACAAATTCGAGAAATGCTCAACACCATC 1209
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
420 GlnValAspGlySerGluLysIleValSerIleLysAspPheLeuG 436
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1210 CAGGTGATGGAGCGAGAAAAAATGCTCGATAAGACATTTCTTGG 1259
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
436 ySerGluAsnLysArgThrGlyAlaLeuGlyAsnLeuLysAsnSerT 453
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1260 AAGTGAGATTAAGAAACCGGGCGTGGGTATCTTAATAAACTCATACT 1309
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
453 erTrpAsnLysAspAsnAsnGluLeuSerHisPheAlaThrThrCys 469
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1310 CTATTAATTAAGATTAATGAATATATCTCATCTTGCACACCTGCTG 1359
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
470 AspLysSerArgProLeuAsnAspLeuValSerGlnLysThrGlnLe 486
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1360 GATAGTCCAGCGCGCTCAACGACTTGTGTAACCAAAAAACAACCTC 1409
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
486 uSerAspIleThrSerArgPheAsnSerAlaIleGluAlaLeuAsn 503
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1410 GTCTGATATTCATCACGTTTAAATTAAGCTATTGAAGCACTGAACCG 1459
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
503 heIleGlnLysThrAspSerValMetGlnArgLeuLeuAspAspThr 519
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1460 TCATTCAAGAAATATGATTCAGTGTGCAACGTCTGCTAGATGACACT 1509
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
520 GlyLys 521
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1510 GGTAAA 1515

```

seq_name: /cgn2_6/ptodata/1/lna/5D_COMB.seq:us-08-913-477-16

seq_documentation_block:

; Sequence 16, Application US/08913477

; Patent No. 5985285

; GENERAL INFORMATION:

; APPLICANT: Tibball, Richard W.

; APPLICANT: Williamson, Ethel D.

; APPLICANT: Leary, Sophie E.C.

; APPLICANT: Bennett, Alice M.

; TITLE OF INVENTION: VACCINES FOR PLAGUE

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHYTE P. C.

Wed Aug 23 11:46:33 2000

us-08-699-716a-2.rni

Page 3

STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ. ID NO. 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1462 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 8..1447
OS-08-913-477-16

```

alignment_scores:
  Quality: 2380.50      Length: 482
  Ratio: 5.001          Gaps: 1
Percent Similarity: 98.755      Percent Identity: 98.133

alignment_block:
US-08-699-716A-2 x US-08-913-477-16 ..

Align seg 1/1 to: US-08-913-477-16 from: 1 to: 1462

43 ASNALAAlaspleuThAlaSerThrhAlaThAlaThleuValGl 59
   :::::|||||
2 AGCTGGCAGATTTAAGTCTGACAGCACCGACTGCAGACGCACTTGTGTGA 51
   |||||
59 UPFOALAArGlIeThrIeuThrTYrLYsGlUyAlaProlIeThrIleW 76
   |||||
52 ACCAGCCCGCATCTCATCATATAAGAGAGCGCTCCAAATTAACATTA 101
   |||||
76 eLsPAsnGLyASnIleAsPhrIguLeuLeuValGlyThrIleuThrIeu 92
   |||||
102 TGGCAATGGCAACATCCATACAGAAATTAAGTGTGTGGTGGCTTACCCT 151
   |||||
93 GLyGLyTYrLYsThrGLyThrThSerThSerValAsnPherThAspAl 109

```

152 GCGGCGTTAAAGAGAGACACCTAGCACATCTTTAACTTTAAGATGC 201
109 aAlGclYasPrOmetYrleuthrPhehrSerGlnAspGlyAsnAsnH 126
202 CGCGGATGCCCATGTACTTAACATTTACTTCTCAGATGGAAATTAAC 251
126 lSgInPhehrThrLysValIlleGlyLysAspSerArgAspPheAspIle 142
252 ACCATTTCACACAAAGATGATGGCAAGATTTCTAGAGATTTTGATATC 301
143 SerProLysValAsnGlyAsnLeuValGlyAspAspValIleLeuAl 159
302 TCTCCTTAAGGTAAACGTGGAACCTTGTGGGGATGACGTGCTTGGC 351
159 aThrGlySerGlnAspPhePheValArgSerIleGlySerLysGlyLyl 176
352 TACGGGCGACGAGATTTCTTGTTCGCTCAATTTGGTTCCAAAGCGGTA 401
176 yLsLeuAlAlaGlyLysTyrThrAspAlaValThrValThrValSerAsn 192
402 AACTTCGACGAGTAAATACACTATGCTGTAAACCGTAAACGTAATCAAC 451
193 Gln.....GluPheMetIleArgAlaTyrGlyGlnAsnProGlnHl 206
452 CAAGATCCATCAGGAGTGTATTAGACCTTACGACAAACCAACACA 501
206 sPheIleGluAspLeuGlyLysValArgValGluGlnLeuthrGlyHsG 223
502 TTTTATTAGAGATCTAGAAAGATTAGGTGGACAACTTACTGTATG 551
223 lYserSerValleuGluGluLeuValGlnLeuValLysAspLysAsnIle 239
552 GTTCTTCAGTTTGTAGAAATTTGGTTCAGTTAGTCAAGATAAAATATA 601
240 AspIleSerIleLysTyrAspProArgLysAspSerGluValIlePheAlaAs 256
602 GATATTTCATTAAATTAAGATCCAGAAAGATTCGGAGGTTTGGCAA 651
256 nArGValIleThrAspAspIleGluLeuLeuLysLleLeuAlaTyrP 273
652 TAGACTATTACTGATGATATCGAATTTCCTCAAGAAATCTACTATT 701
273 heLeuProGluAspThrIleLeuLysGlyHsTyrAspAsnGlnLeu 289
702 TTCTACCGAGAGATCCATCTTAAAGCGGTCATATTAGACAACCAACTG 751
290 GlnAsnGlyIleLysArgValLysGluPheLeuGluSerProAsnTh 306
751 CAATAATGCATCAACCGAGTAAAGAGTCCCTTGATCATATCGCGAATAC 801
306 rGlnrProGluLeuArgAlaPheMetAlaValMetHisPheSerLeuthrA 323
802 ACAATGGGATTTGGCGGCTTCATATGCGAGTAATGATTTCTTTAACCG 851
323 lAspAspArgIleAspAspAspIleLeuLysValIleValAspSerMetAsn 339
852 CGGATCGATATGATGATATTTTGAAAGAGATTTGATTCAATCATGAT 901
340 HisHisGlyAspAlaArgSerLysLeuArgGluLeuAlaGluLeuth 356
902 CATCATGTGTGATGCCCGGTAGCAAGTTGCGTCAAGAAATAGCTGAGCTTAC 951
356 rAlaGluLeuLysIleTyrSerValIleGlnAlaGluIleAsnLysHisL 373
952 GCGCGAATTTAAAGATTATTCAGTATTCAAGCGGAAATTAATAAGCATC 1001
373 euSerSerSerGlyThrIleAsnIleHisAspLysSerIleAsnLeuMet 389
1002 TGTCTAGTAGTGGCACCAATAATATCCATGATTAATCATTAATCTCATG 1051
390 AspLysAsnLeuTyrGlyTyrThrAspGluGluIlePheLysAlaSerAl 406

seq=documentation_block:
: Sequence 1, Application US/08913477
: Patent No. 5985285
: GENERAL INFORMATION:
: APPLICANT: Tibball, Richard W.
: APPLICANT: Williamson, Ethel D.
: APPLICANT: Leary, Sophie E.C.
: APPLICANT: Oyston, Peter C.F.
: APPLICANT: Bennett, Alice M.
: TITLE OF INVENTION: VACCINES FOR PLAGUE
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: NIXON & VANDERHAYE P.C.
: STREET: 1100 No. 5985285th glene Rd. 8th floor
: City: Arlington
: STATE: VA
: COUNTRY: USA
: Zip: 22201-4741
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/913.477
: FILING DATE: 15-SEP-1997
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/GB96/00571
: FILING DATE: 13-MAR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9505059.7
: FILING DATE: 13-MAR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9518946.0
: FILING DATE: 15-SEP-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9524825.8
: FILING DATE: 05-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.

194	GLPHEMERILEYSLAYGLINDINSNPROGINSPEHIEGLIAS	210	PLEUGLULYSVALARGVALGLUGLLEUTHGLYHISGLYSESERVAL
195	GLPHEMERILEYSLAYGLINDINSNPROGINSPEHIEGLIAS	211	PLEUGLULYSVALARGVALGLUGLLEUTHGLYHISGLYSESERVAL
7	GAATTC...ATRAGACCTCAGAACAAAACCCACACACTTTATTGAGA	53	
227	EUGLUGLLEUVALGLINEUVALYLSAAPLYSASNILEASPLIESERILE	244	
104	TAAAGAAATTGGTTCAGTTAGTACGAAGAATAAAAATATGATATTTCCATT	154	
244	LYETIYASPRCAATGLYSASPERGLUVALPHEALASARGVALILETH	260	
154	AAATATGATCCCGAAGAAAGATTGGAGGTTTTGGCAATACAGTAATTAC	204	
260	TAAPASPIELGLLEULEUVALYSLILEUVALATRYPHELEUPROGLUA	277	
204	TGATGATATCGAATTCCTCAAGAAATCCATGCTAATTTTCTACCCGAGG	294	
254	ATGCCATTTCTTAAGCGGCTCATATGACAAACCAATGCAAAATAGGCATC	304	
294	LYEATGVALYSGLPHELEUINSERSEPRASINTRINTRIGLUE	310	
304	AACGAGTAAAGAGTCTCTTAATCAATCGCGAATACACATGGCAATT	327	
310	UARGALAPHEMETALVALMETHISPHESERLEUTHRALASPRATLEA	334	
354	GCGGCGGTTCAAGCGCAGTATCATTTCTTTAAACCCGCAATCGTATCG	404	
327	SPASAPSPILEUULYSVALILEVALASERMETASNHISHSGLYASP	444	
404	ATATATATATTTTGAAGAGGATGTGATTAACAAACATCAACAGGGGAT	454	
344	ALAEATSERLYLEUARGLUGLLEUVALAGLUEUTHRALAGLUEULY	504	
454	GCCCGTAGCAAGTGGCGTAAACAAATTAACCTAGCTACCGCCGCAATTTAA	560	
560	SILETYRSEVALILEGINALAGLULEASNLYSLEUSERSEBERG	3777	


```

330 leuLeuysValIleValAspSerMetCasnHisHISglYsPheIaaYsSer 346
347 LysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyrSe 363
463 AAATGGCGTGAAGAATTAGCTGAGGCTTACCGCGAATTAAAGATTATTC 512
363 rValIleGlnAlaGluIleAsnLysHisLeuSerSerSerglyThrIleA 380
513 AGTTATTCAAGCCCAAAATTAAATACCATCTGCTGTAGTGGCACATTA 562
380 snIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGlyTyr 396
563 ATATCCATGATTAATTCATTATATCTCAGATGATAAAATTTATATGGTTAT 612
397 ThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuLysI 413
613 ACAGATGAAGAAGATTTTAAAGCCGACGCCAGAGTACAAATTCCTCAGAA 662
413 sMetProGlnThrThrIleGlnValAspGlySergLysLysIleValS 430
663 AATGGCCCAAAACCCACCATTCAGGTGAGTGGAGCGCAGAAAAAAATAGTCT 712
430 erIleLysAspPheLeuGlySergLysAsnLysArgThrGlyAlaLeuGly 446
713 CGATTAAGGACTTCTTGGAAAGTGGAATTAAGAAACCGGGCGTTGGGT 762
447 AsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSerH 463
763 AATCTGAAAAACATCATCTTATTATATAAGATTAATTAATGAATTATCTCA 812
463 sPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuValS 480
813 CTTTGGCACACCTGCTCGATTAAGTCCAGCGCGCTCAACGACTTGTTA 862
480 erGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSerAla 496
863 GCCAAAAAACAACCTCAGCTGCTGTATTATTAATCACTGCTTTAATTAGCT 912
497 IleGluLalaLeuAsnArgPheIleGlnLysTyrAspSerValMetGlnAr 513
913 ATTTGAGCACTGAACCGTTTCAATTCAGAAATATGATTCACTGATGCAACG 962
513 glLeuLysAspAspThrSergLysLys 521
963 TCTGCTGATGACACGCTGCTGTA 967
seq_name: /cgn2.6/ptodata/1/ina/5D_COMB_seq:US-08-913-477-20

```

seq.docdocumentatn_block:
Sequence 20. Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Peta C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHART P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA
 APPLICATION NUMBER: US/08/913,477
 FILING DATE: 15-SEP-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/GB96/00571
 FILING DATE: 13-MAR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9505059.7
 FILING DATE: 13-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9518946.0
 FILING DATE: 15-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9524825.8
 FILING DATE: 05-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Crawford, Arthur R.
 REGISTRATION NUMBER: 25,327
 REFERENCE/DOCKET NUMBER: 124-599
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-816-4000
 TELEFAX: 703-816-4100
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 547 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Yersinia pestis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 29..538
 US-08-913-477-20

```

alignment_scores:      Quality: 846.00      Length: 171
                       Ratio: 4.947      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 98.830

alignment_block:
US-08-699-716A-2 x US-08-913-477-20  ..

Align seg 1/1  to: US-08-913-477-20  from: 1  to: 547

23  HSMetLysLysLLeSerSeValIleAlaIleAlaLeuPheGlyThrI  39
:::|||||
26  TTAAGAAGAAAAATAGTCCGTTATGCCATTGCATATTGGAACTAT  75
39  eAlaThrAlaAsnAlaAlaAspLeuThrAlaSerThrThAlaThrAla  56
76  TCCAACTGCTATGGCGGAGATTTAACTGCAGACACACATCCAAAGCAA  125
56  hLeuValGluProAlaArgIleThrLeuThrTyrLysGluGlyAlaPro  72
126  CTTGTGGTGAACCAAGCCCGCATCTCATCATTAAGGAAGCGGCTCCA  175
73  IleThrIleMetAspAsnGlyAsnIleAspThrIleuLeuValGlyTh  89
176  ATTACAAATTAAAGCAAAAGGAACATGCATACAAATAATTACTGTGGTAC  225
89  rLeuThrIleuGlyGlyTyrLysThrGlyThrThSerThrSeValAsp  106
226  GCTTACTCTTGCGCGCTTAATAACAGGACACATGACATCTGTACT  275
106  heThrAspAlaAlaGlyAspProMetGlyrLeuThrPheThrSerGlnAsp  1222
276  TTACAGATGCCCGGGGTATCCCATGTACTTAACTTACATTCTTCAGAT  325

```

```

123 GYASnAsnHISGlnPheThrThrLysValIleGlyLysAspSerArgAs 139
|||||
326 GGAATAACCAATTCACATCAAAAGTAGTGGCAAGATTCCTAGAGA 375
|||||
139 pPheAspIleSerProLysValAsnGlyLysAsnLeuValGlyAspAsp 136
|||||
376 TTTTGATATCTCTCCCTAGGTAACGCTGAGAACCTGTGGGGATCAGG 425
|||||
156 aValLeuAlaThrGlySerGlnAspPhePheValArgSerIleGlySer 172
|||||
426 TCCTCTTGGCTACGGCAGCAGCATTTCTTTCTTCCTCAATTGGTCC 475
|||||
173 LysGlyGlyLysLeuAlaIleGlyLysThrAspAlaValThrValTh 189
|||||
476 AAGCGCGGTAACTGACAGCAGTAATACATGATGCTGTAACCGTAAC 525
|||||
189 rValSerAsnGln 193
|||||
526 CGTATCTAACCA 538

seq_name: /cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-913-477-10

seq_documentation_block:
; Sequence 10, Application US/08913477
; Patent No. 5985285
; GENERAL INFORMATION:
; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Leary, Sophie E.C.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Bennett, Alice M.
; TITLE OF INVENTION: VACCINES FOR PLAQUE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYTE P.C.
; STREET: 1100 No. 5985285th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913.477
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00571
; FILING DATE: 13-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9505059.7
; FILING DATE: 13-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9518946.0
; FILING DATE: 15-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9524825.8
; FILING DATE: 05-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 124-599
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Yersinia pestis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..454
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: 1..6
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: 536..541
; US-08-913-477-10

alignment_scores:
Quality: 755.00 Length: 151
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.013

alignment_block:
US-08-699-716a-2 x US-08-913-477-10 ..

Align seg 1/1 to: US-08-913-477-10 from: 1 to: 541

43 AsnAlaIleAspLeuThrAlaSerThrThrAlaThrAlaThrLeuValG1 59
|||||
2 AGCTCGCAGATTACTGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 51
|||||
59 uproAlaArgIleThrLeuThrThrLysGlyLysAlaProIleThrIle 76
|||||
52 ACCGCGCCGATCCTATTTACATATTAAGAGAGCGCTCCAAATTACATTA 101
|||||
76 eAspAsnGlyAsnIleAspThrGluLeuLeuValGlyThrLeuThrLeu 92
|||||
102 TGGACATAGGAACATCGATACGAATTTCTTGTGGTACGCTTACTCTT 151
|||||
93 GlyGlyThrLysThrGlyThrThrSerThrSerValAsnPheThrAsp1 109
|||||
152 GGGGCTATATAAACAAGACACACTACACACTCTGTTAACTTACAGATC 201
|||||
109 AlaGlyAspProMetThrLeuThrPheThrSerGlnAspGlyAsnAsn 126
|||||
202 CGCGGCTGATCCCATGTACTTAACTTACTTCTCGAGATGGAATAAC 251
|||||
126 IsGlnPheThrThrLysValIleGlyLysAspSerArgAspPheAsp1 142
|||||
252 ACCAATTCACCTCAAAAGTAGTGGCAAGATTCTAGAGATTTGATATC 301
|||||
143 SerProLysValAsnGlyLysAsnLeuValGlyAspAspValAlaLeu1 159
|||||
302 TCCTCTTAAGTAACCGTGAACCTTGGGGGAGAGAGAGCTGCTTGGC 351
|||||
159 aThrGlySerGlnAspPhePheValArgSerIleGlySerLysGlyL 176
|||||
352 TAGCGGCACCCAGAGATTCTTGTGCTCAATTTGTTCCAAAGCGGTA 401
|||||
176 yLeuValAlaIleGlyLysThrThrAspAlaValThrValThrValSer 192
|||||
402 AACTTCAGCAGGTAAATACATGATGCTGTAAACCTAACCTATCTAAC 451
|||||
193 Gln 193
|||
452 CAA 454

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-792-865D-1
seq_documentation_block:
; Sequence 1, Application US/07792865D

```

Patent No. 5546247
GENERAL INFORMATION:
APPLICANT: John W. Birnwell, Mary W. Galinski,
APPLICANT: Samuel P. Wertheimer
TITLE OF INVENTION: MEROZOITE ANTIGENS LOCALIZED AT
TITLE OF INVENTION: THE APICAL END OF THE PARASITE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dady & Dady P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
COMPUTER: IBM or IBM-compatible
OPERATING SYSTEM: PC/MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/792,865D
FILING DATE: 19911004
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US90/01849
FILING DATE: April 3, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, Arthur
REGISTRATION NUMBER: 34,354
REFERENCE/DOCKET NUMBER: 5986/14692-US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3763 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
DESCRIPTION:
HYPOTHEetical: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: P.vivax
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY: Lambda gt 11 native P.vivax
LIBRARY: DNA expression library
CLONE: 5.3
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: This sequence corresponds to
OTHER INFORMATION: Figure 1A (sheets 1-4) in the
OTHER INFORMATION: application, as filed.
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:

```

: VOLUME:
: ISSUE:
: PAGES:
: DATE:
: DOCUMENT NUMBER:
: FILING DATE:
: PUBLICATION DATE:
: RELEVANT RESIDUES IN SEQ ID NO:
US-07-792-865D-1

alignment_scores:
      Quality: 164.50      Length: 645
      Ratio: 0.532      Gaps: 33
Percent Similarity: 47.907      Percent Identity: 20.310

alignment_block:
US-08-699-716A-2 x US-07-792-865D-1 ..

Align seg 1/1 to: US-07-792-865D-1 from: 1 to: 3763

      8 HISHISHISHISerSerGlyHISleAspSpsasp..... 21
      |||.....:|||||
1201 CATAAATTATACACGCGAAGATGAAGACTTAATATATATATACGAA 125C
22 .....LysHisMetLysL 26
1251 TCGAAGGCTACTATAGAACTACCTAGAAAGCTAAACAT.....A 1294
26 ysIleSerSerValIleAlaIleAlaLeuPheGlyThrIleAlaThrAla 42
      ::|||::|||::|||
1295 ATTATACAGAAATTAATCTTAATATTAAACGAGGAGAGAAAAATATACAGT 1344A
43 AsnAlaIleAspLeu.....ThrIleSerThrThrAl 53
      ::|||::|||::|||
1345 AAAGCTTAAGATATCATGCAGAAAAATTAAGACACTTCAGAAATACTGC 1394A
53 arThrAlaThrLeu..... 57
      | |||||
1395 AGAGAAACCTTGTAGAGAGGTGAAGACGACCATCTAATATATGTTAAT 1444A
58 .....ValGluProAla 61
1445 ATTAAATCAAAATACCCACAGAAAGAAATCTTATCGTTACGCAAAAAAT 1494A
62 ArgIle.....ThrLeuThrTyrLysGluGlyAlaProI 73
      |||::|||::|||
1495 AGCATTAATGTGTAGATTCCTCACTATTACAAATATAGAAAGGCGA.... 1539S
73 eThrIleMetAspAsnGlyAsnIleAspThrGluLeuLeuValGlyThrL 90
      ::|||::|||::|||
1540 .CTTAAGATCCAGAGGAATATGAAATGCAATTTTGGAAAAAGTTAG 1588S
90 eutThrLeuGly.....GlyTyrLysThrGlyThrThrSer 103
      ::|||::|||::|||
1589 AAGCAATAGCTAAAAATAGAAAAATTAAAGGTGACATNACCAAAAAATCA 1638S
104 ValAsnPheThrAspAlaIleGlyAspProMetTyrLeuThrPheThrSe 120
      ::|||::|||::|||
1639 ATTAATTTCAACATGCGGAACCTTTCTCCCTCTTCAACAACTTT.... 1683S
120 rGlnAspGlyAsnAsnHisGlnPheThrThrLysValIleGlyLysAsps 137
1684 .....GAT 1687
137 eArGAspPheAspLleSerProLysValAsnGly...GluAsnLeuVal 152
      ::|||::|||::|||
1688 TAAATCAATGTGCTTAAATTAATAATTAATGATTTTGAAAAATTAATG 1737S
153 GlyAspAspValValLeuAlaThrGlySerGlnAspPhePheValArgSs 169
      |||::|||::|||
1738 GCGGAATATATATACGAATTTGGAAGGATCTTAAATTAATTAATTTAGTGA 1787S

```

```

169 rllleglyserlysglyglyleualaalaaglylytyrthraspallav 186
      ::::::::::::::::::::
1788 tttta.....AGAAATGCTTCGAAAAACCTTCAGACTATTAACAGCAA 1831
      ::::::::::::::::::::
186 althrvalthrvalserasnnglnluphemetlearglatyr..... 200
      ::::::::::::::::::::
1832 AACACGTGAGCGCTAGAGCAGCAAAAAAGTTAATCTATTAATAATAA 1881
      ::::::::::::::::::::
201 glnlgnasnprglnhispheileglnaspheuglnlyval..... 214
      ::::::::::::::::::::
1882 GAAGAAGAGCAAAATATAATTTAAGAGATGTTAAAAAGTGAATCATT 1931
      ::::::::::::::::::::
214 ..... 214
1932 CAGATTATATTTAATATGAAAGAAAGCTTAGATTATAGATGA 1981
      ::::::::::::::::::::
215 ..Argvalgluglnleuthr.....Glyhisgllyser.....Ser 225
      ::::::::::::::::::::
1982 TTTAAAAAAGAACAACTACAGTCAATGAAGACGCTAACGTTAAACAA 2031
      ::::::::::::::::::::
226 valleuglnlgluleuvalgluleuvallyspasplyasnleapilese 242
      ::::::::::::::::::::
2032 CTAGTGGAAATTTAAAGGTTAGTTGATGAACAACACTTATCAGATAT 2081
      ::::::::::::::::::::
242 rllelystyrspproarglylsaspsergluvalphenalaasnargvali 239
      ::::::::::::::::::::
2082 ATTAAAAAGACGACGCGCAAAAATAGAGAAATA..... 2115
      ::::::::::::::::::::
259 lethraspaspilleglnleuleuvallystleleualaryrphelupro 275
      ::::::::::::::::::::
2116 ..... 2116
276 glu.....Aspthrilleuvallysglyglyhislytyraspsnnglnleugl 290
      ::::::::::::::::::::
2143 AATAAAGCAAAACCTATTTTA.....GACACGTAGATCTCTGCCAA 2186
      ::::::::::::::::::::
290 nasn..Glyllelysarvallysgluphe..... 299
      ::::::::::::::::::::
2187 ATATGTAGCATTAATAAACCTGAGTTGCGACATAACAGAAATGTTAG 2236
      ::::::::::::::::::::
300 .....leugluserSerprobsnthrglntrpgluleuarg... 311
      ::::::::::::::::::::
2237 GAGATGCAAAATTTGAAACTGACAGCAAGAAATAATTGTCGCAAAAAAT 2286
      ::::::::::::::::::::
312 ...AlaphemetalayalMethispheserleuthralasparglyleas 327
      ::::::::::::::::::::
2287 AATGTGTACTAGAAACAGAAAAATATGTCAAGAATTCACAAACGAAATGGA 2236
      ::::::::::::::::::::
327 P.....Aspaspilleuvallyvallevalaspsersmeta 339
      ::::::::::::::::::::
2337 TGTTCATTAATAAATATACAGATGCTTACAAAGTTGCGACGTGAATACCTG 2386
      ::::::::::::::::::::
339 snhishtisgly.....Aspalargserlystleuarglgluleuvala 353
      ::::::::::::::::::::
2387 CCCACCTCAGACGAATATAGATACCAAAACAAAGACAGTTCTAATAATAA 2436
      ::::::::::::::::::::
354 glnleuthralagluleuvallystlyrservalilleglnalaglulleas 370
      ::::::::::::::::::::
2437 GAAATGGGAACCAAAATATCTTAAAGTTGTCTA.....ATAAA 2477
      ::::::::::::::::::::
370 nlyshstleuserSersegllythrlleasnlehisaspserlyleas 387
      ::::::::::::::::::::
2478 TCATATCAAAAATAAATAAGCTCTAATAAAGTAAGGAAGACGCTTT 2527
      ::::::::::::::::::::
387 snleuwtasplyasnleuvaltyrtyrthraspgluglullepelys 403
      ::::::::::::::::::::
2528 CAGTCAAAATAGGTAATGTT...TCCAAAGAACATAGTGGTTAAGCAAA 2574
      ::::::::::::::::::::
404 AlaserAlaglutyrllys.....lleuglulysme 414
      ::::::::::::::::::::
2575 ATTACATGCGCGATAAAGTTACGATACATACGCTTAGAGAAA.. 2622
      ::::::::::::::::::::
414 tproglnthrthrlleagl..... 420

```

```

2623 ..... 2668
2623 ...CAAATGATTAACAAAATCTACGCAATCTTTCACCTACAGAAAAGA 2668
      ::::::::::::::::::::
421 ..Valaspilyserglulyslystlevalserllelys...Aspbeleu 435
      ::::::::::::::::::::
2669 CTAAACAGAAATGCGATTCGAAAGTTGGAATAAATTAACAGATTTCCGA 2718
      ::::::::::::::::::::
436 glysergluasnlysarthrvalaleugllyasnleuvalys..... 449
      ::::::::::::::::::::
2719 AATTGCAAAAT.....GCATTAACAAACACTGAAGAGAACT 2756
      ::::::::::::::::::::
450 .Asnserlyserlyrasnlyspasnnglnleuserhisphenalat 466
      ::::::::::::::::::::
2757 AATGCTCTAAAGCAAGCTCGACATCATGAA..... 2790
      ::::::::::::::::::::
466 hrthrcysserasplysserargproleuasnaspheulvalserglnlys 482
      ::::::::::::::::::::
2791 ..CANGTACAAAGTAAGTAAGTGAACCACTAATCTGCGCTATCCGAAAT 2838
      ::::::::::::::::::::
483 thrthrglnleuseraspilletterserargpheasnseralileglual 499
      ::::::::::::::::::::
2839 GAAAAAGAAAGAACGACATACATAGT...CTTATACGCGCCCTTGATGA 2885
      ::::::::::::::::::::
499 aleu.....Asnargpheileglnlyst 507
      ::::::::::::::::::::
2886 GTTATTAATAAAGAACAGACATCGCAAGTACTAGTACAAACTGATTA 2935
      ::::::::::::::::::::
507 yraspservalmetglnargleuleuaspaspthr 518
      ::::::::::::::::::::
2936 AGGATACCGTTACCAAGAAATAGTGTAGTACAC 2970
      ::::::::::::::::::::
seq_name: /cgn2_6/ptodate/1/lna/5B_COMB.seq:us-08-480-604A-25
seq_documentation_block:
; Sequence 25, Application US/08480604A
; Patent No. 5736139
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.
; APPLICANT: PADHYE, NISHA V.
; APPLICANT: FIRCH, JOSEPH R.
; APPLICANT: STAFFORD, DOUGLAS C.
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,604A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/422,711
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/405,496
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993

```



```

290 nasnllylleysarvalylsgluipheleuglSerSeProsrnhrng 307
    |||
653 TAAC.....ATCCAGCCTTCAATTAACATCA 658
307 intrpGlueu.....ArgalapheMetalaValmet 317
    :|||:||||
659 TGTTCMAACGTGACGGTGTGCTGACACTACACCGCTACATCTGATCAAA 708
318 HtSPheSerLeuthrAlaAspArgTlleAspAspArglleuylVal11 334
    :|||:|||||:|||||:|||||:|||||:
709 TACTTCATCTGTTTCGACAAAGAACTGACCAAAAGAAATCAAAAGACT 758
334 evalAspSerMetasnHtSHtSGlAspAlaArgSerlyLeuArg1ug 351
    :|||:|||||:|||||:|||||:|||||:
759 GTACGACACCAAGTCCAACTTGTGT.....ATCCTGAAGAGCT 796
351 lulevalGlueleuthrAlaGlueuylSerVal1leglnAla 367
    :|||:|||||:|||||:|||||:|||||:
797 TCTGGGGTACTACTCTGACAGTACGACAAACCGTACTACATG..... 837
368 GluIleasnlySHtLeuSerSerSerglyThrIleasnIleHtasp.. 383
    :|||:|||||:|||||:|||||:|||||:
838 .....CTGAATCTGTACGATCC 854
384 ... lysSerIleasnleuMetaplysAsnleuTyrglyThrThrspg 399
    |||
855 GACCAAAATACCTGTACCGTCCGACAAATGATGATATCCGGGTTAC..... 897
399 luguIlePheLyAlaSerAlaGluTyrlsIleuGluTysMetPro 415
    .....
897 ..... 897
416 GlnthrThrIleglnValAspGlySerGluTysIleValSerIlely 432
    :|||:|||||:|||||:|||||:|||||:
898 .....ATGTACCTGAAGGTCGCGGTCTGTCTGTATGACTACCAA 938
432 sasPheleuulysSerGluAsnlySArgThylAlaLeuGlyAsnleul 449
    :|||:|||||:|||||:|||||:|||||:
939 CATCTACCTGACACTTCTCCCTGTACTGC...GGTACCAAAATCATCATCA 985
449 ysAsnSerTySerTyAsnlyAspAsn 458
    |||:|||||
986 AGAAATACGCTGTGTTACCAAGACAACT 1014

seq_name: /csgn2.6/ptodata/1/lna/5_COMB.seq:US-08-870-370-7
seq_documentation block:
  Sequence 7, Application US/08870370
  Patent No. 6060242
  GENERAL INFORMATION:
  APPLICANT: Eileen Nie and Yuan Min Wu
  TITLE OF INVENTION: PNA Diagnostic Methods
  NUMBER OF SEQUENCES: 13
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Caesar, Rivise, Bernstein, Cohen &
  ADDRESSEE: Pokotilow, Ltd.
  STREET: 12th Floor, 7 Penn Center, 1635 Market
  STREET: Street
  CITY: Philadelphia
  STATE: Pennsylvania
  COUNTRY: U.S.A.
  ZIP: 19103-2212
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/870,370
  FILING DATE:
  ATTORNEY/AGENT INFORMATION:
  NAME: Tenet, David M.
  REGISTRATION NUMBER: 37,054

```

```

1 REFERENCE/DOCKET NUMBER: E1047/20001
2
3 TELECOMMUNICATION INFORMATION:
4
5 TELEPHONE: 215-567-2010
6
7 TELEFAX: 215-751-1142
8
9 INFORMATION FOR SEQ ID NO: 7:
10
11 SEQUENCE CHARACTERISTICS:
12
13     LENGTH: 375 bases
14
15     TYPE: nucleotide
16
17 STRANDEDNESS: double-stranded
18
19 TOPOLOGY: linear
20
21 US-08-870-370-7

```

```
alignment_scores:      length: 22
                      quality: 146.00
                      ratio: 6.636
                      gaps: 0
Percent Similarity: 100.000
Percent Identity: 100.000
```

```
alignment_block:
US-08-699-716A-2 x US-08-870-370-7 ..
```

Align seg 1/1 to: US-08-870-370-7 from: 1 to: 375

17 eapapapapapaply 22
139 CGAGCGAGCGAGCGAG 154

seq_name: /cgn2_6/ptodata/1/lna/6_COMB.seq:US-08-870-370-8

```

seq.documentation block:
Sequence 8, Application US/06870370
Patent No. 6060242
GENERAL INFORMATION:
APPLICANT: Eileen Nle and Yuan Min Wu
TITLE OF INVENTION: PNA Diagnostic Methods
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Caesar, Rivise, Bernstein, Cohen &
ADDRESSEE: Pokocilow, Ltd.
STREET: 12th Floor, 7 Penn Center, 1635 Market
Street
City: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103-3212
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/06/870,370
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Tenet, David M.
REGISTRATION NUMBER: 37,054
REFERENCE/DOCKET NUMBER: E1047/20001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-567-2010
TELEFAX: 215-751-1142
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 bases
TYPE: nucleotide
STRANDEDNESS: double-stranded
TOPOLOGY: linear
IS-08-870-370-8

```



```

1368 TGCAGGCTGCTTAAAGAAAGTGCCTACTAAATGCAAC..... 1411
108 AspaLaalaglyAspPrometYrLeuThrPheThrSerGlnaspGlyAs 124
1412 .....GATCTTCATCAAGCTCTCACAC 1434
124 nasnHisGlnPheThrThrLysValIleGlyLysAspSerArgasp... 139
1435 CGTCTTACATCAACAAACATAGTTGAGCAAGATGCAAGACACATG 1484
140 .....PheAspIleSerProLysValasnGlyLysasnLeu 151
1485 GACAAACAAATAATTTGTTAAATGGCAATGGAATTTGCTGATTTG 1534
152 ValGlyAspAspValValLeuAlaThrGlySerGlnaspPheValar 168
1535 GCAGCGAGTGAACATCAACAGATCGGCTCGGAGATAAAGGCGTCA 1584
168 gserIleGly...SerLysGlyLysLeuAlaAlaGlyLysThr 184
1585 AGAAGCTGGCTTAATAAACAATCGCTGTAACACTAGCGCTGTATCA 1634
184 spAlaValThrValThrValSerasnGlnIuphMetIleArgAlaTyr 200
1635 ACGCAGTC...GTGATCATCTTCAACAT.....ATACCTTACAGA 1672
201 GlnGlnAsnProGlnHisPheIleGlnasp.....LeuGlyLys 213
1673 GAATCTAAGCTAACACAGATGCTCAAGACTCTTAGGTGATGACGAA 1722
213 svaIArgValGlnGlnLeuThrGlyHisGlySerSerValLeuGlnul 230
1723 AACATGCAATTATGCCAACTATATCACTGCGAAATATTCATGGAAGA 1772
230 euValGlnLeuValLysAspLysasnIleAspIleSerIleLysTyrAsp 246
1773 CTGCAACTAGCTGATGATGATGCAACGAGACCAATCATTAAGAAACT 1822
247 ProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAsp 263
1823 CCACAGTAATCAATCACTCT..... 1840
263 eglLeuLeuLysLysIleLeuAlaTyrPheLeuProGlnAspThrIle 280
1841 .....TTATCGAAGATATCACTGTC 1859
280 euLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArgVal 296
1860 TCAAA.....GACTACATTCAGAG..ATTGAAATA 1891
297 LysGlnPheLeuGlnSerSerProAsnThrGlnTrpGlnLeuArgAlaPh 313
1892 AGAATGATTTGAAAAATTCAGAAACAACA..... 1924
313 emeAlaValMetHisPheSerLeuThrAlaAspArgIleAsp...Asp 329
1925 .....GCTATATTATACCAAGATCACTGAGCTTACG 1961
329 spIleLeuLysValIleValAspSerMetAsn.....HisHisGly 342
1962 AGACCAATTCATCTGATTCATGACCAAAATCTAAAAATCATATAC... 2008
343 AspaLaalagSerLysLeuArgGlnGlnLeuAlaGlnLeuThrAla.... 357
2009 .....CTGCGAGAACAAATTAATAAATTCAGAAAGAACTA 2043
358 .....GlnLeuLysIleTyrSerValIleGlnAlaGlnLeuLysn 372
2044 CCGAACCAATTCATGATATCAATAATCTTTTACAGTCAAGAAAAAGAAAC 2093
372 LysLeuSerSerSerGlyThrIleAsn..... 380
2094 TAATTCGCAATATACAGAAATTTAATGTCATTTTCTACTTTACTCG 2143

```

```

381 .....IleHisAspLysSerIleAsnLeuMetAspLysAsnLe 393
2144 GAATCCAAAAAATTCACACTACTAATCTGCAACTGAAATGA...GT 2190
393 uTyrGlyTyrThrAspGlnGlnIlePheLysAlaSerAlaGlnTyrLys 410
2191 CATACAAACAGAGATTTTTCACCTAGAAATTTCTCAAAAACAGTATATA 2240
410 LeuGlnLysMetProGlnThrThrIleGlnValAspGlySerGlyLys 426
2241 CGAACCGAAGCATG.....CAATTAATAATCTCAACA 2275
427 LysIleValSerIleLysAspPheLeuGlySerGlnasn..... 439
2276 GTTTACAGACTTTGAAACACTTACAGGCTCTTAAATATATATACCTC 2325
440 LysArgThrGlyAlaLeuGlnLysasnLeuLysAsnSerTyrSerTyrAsn 456
2326 TAATGTTCGAAAGTTATCAAGGCGTCAACGAAAGATCAACAGGAACG 2375
456 LysAspAsnAsnGlnLeuSerHisPheAlaThr..... 466
2376 TAATACCAATTAAGCGCAACACAGATTCCTCAATGCTATTAAC 2425
467 ...ThrcysSerAspLysSerArgProLeuAsnAspLeuValSerGln 482
2426 ATTTACTACTACTATGATGATGATGATGATGATGATGATGATGAT 2475
482 sThrThrGlnLeu.....SerAspIleThrSerArgPheAsn 495
2476 TTCACCTCTCAATGAAATATTCAGAGTGAATCTACTCTCACTATCGT 2524
495 eraAlaIleGlnAlaLeuAsnArgPheIleGlnLysTyrAspSerValMet 511
2525 .....AAGATTTCATGAATTCACCAATCACTACATCACTACATTCCTA 2566
512 GlnArgLeuLeuAspAsp 517
2567 AAAAATTTACAAAACGAT 2584

seq_name: /cgn2_6/ptodata/1/lna/5C_COMB.seq:US-08-376-843-23
seq_documentation_block:
; Sequence 23, Application US/08376843
; Patent No. 5846764
; GENERAL INFORMATION:
; APPLICANT: Demaggio, Anthony J.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/376,843
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5846764and, Greta E.

```

REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/31784
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5093 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-376-843-23

alignment_scores:

Quality: 129.00 Length: 506
 Ratio: 0.500 Gaps: 22
 Percent Similarity: 50.988 Percent Identity: 19.170

alignment_block:

US-08-699-716a-2 x US-08-376-843-23

Align seg 1/1 to: US-08-376-843-23 from: 1 to: 5093

```

73  ILeThrIleMetAspAsnGlyValIleAspThrGluLeuVal..... 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1268 ATTGCTTTTGGACATACACAAATATTCATCATGCTGCAAGG 1317
88  ....
1318 GATGACGAATCTTATTAAGTCTGACACAGAGCTGATTTGCTAA 1367
91  hIreUGlyGlyTyrIleThrGlyThrSerThrSerValAspPheThr 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1368 TGCAGGCTTGTAAAGAAAGAGAGTGCCTACTAAATGCAC..... 1411
108 AspAlaIleGlyAspPrometTyrLeuThrPheThrSerGlnAspGlyAs 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1412 ..... GATCTTTCATCAAGGCTTCACAC 1434
124 nasnHISglnPheThrThrIleValIleGlyIleAspSerArgAsp.... 139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1435 CGTCTTACATCACAACAACAGATGTTGACAGATGACCAACACCATG 1484
140  .... PheAspIleSerProIysValAsnGlyGluAsnLeu 151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1485 GACAAACAAATAATTTGTAAATTTGGCAATGAAATTTGGTGAATTTG 1534
152 ValGlyAspAspValValLeuAlaThrGlySerGlnAspPhePheValAr 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1535 GCAGCGCTGAACATCAACAGATCGGGTGGGAGATTAAGGGCTCA 1584
168 gSerIleGly...SerIysGlyGlyIleValAlaGlyIleTyrThrA 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1585 AGAAGCTGGCTATATAACAATCGCTACACAGTACGCGGTATATCA 1634
184 sPAlaValIleThrValIleSerAsnGlnIlePheMetIleArgAlaTyr 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1635 ACGGACATC...GTTCATCATTCCTAACCAT.....ATACCTTACAGA 1672
201 GluGlnAsnProGlnHisPheIleGluAsp.....LeuGlyIle 213
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1673 GAATCTAAGCTAACAGATTCCTACAGACTCTTAGCTGATGAGCAAA 1722
213 sValArgValGluGlnIleuThrGlyHisGlySerSerValLeuGluGlu 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1723 AACATGCAATATTCGCAACTATATCACTCGCAAAATATCCATGAGA 1772
230 euValGlnLeuValIleAspIleAsnIleAspIleSerIleIleTyrAsp 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1773 CTGCAAGTACGCTAGAAATATGCAACGAGAGCAAAATCAATTAAGA 1822
  
```

```

247 ProArgIleAspSerGluValPheAlaAsnArgValIleThrAspAspIle 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1823 CCACAGTAATAGTCT..... 1840
263 eGluLeuLeuIleValIleLeuAlaTyrPheLeuProGluAspThrIle 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1841 .....TTATCGAAGATACATGTC 1859
280 euIleGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleIleAspVal 296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1860 TCATA.....GACTACATTCAGAG...ATGAAATATTA 1891
297 IysGluPheLeuGluSerSerProAsnThrGlnIlePheLeuArgAlaP 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1892 AGAATGATTTGAAAATTCAGAAACAAACAA..... 1924
313 eMetAlaValMetHisPheSerLeuThrAlaAspArgIleAsp...Asp 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1925 .....GCTATATTATTAACTCAAGTCAAGTCAAGTCTTACG 1961
329 sPileLeuIleValIleValAspSerMetAsn.....HisHisGly 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1962 AGAGCAATCTATCTTGATTTGATGAGCAAAATCTAAATATACATAC... 2008
343 AspAlaArgSerIleuArgGluGluLeuAlaGluLeuThrAla..... 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2009 .....CTGCGAGAACAAATTAATAATTCAGAAACAACTA 2043
358 .....GluLeuIleIleTyrSerValIleGlnAlaGluIleAsnIle 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2044 CTGAAACCAATTCATATCATATCTTTACAGTCTGAAAGAAAGAAAC 2093
372 ILeuSerSerSerGlyThrIleAsn..... 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2094 TAATTCCTAATATACAGAAATTTAAATGCGATTTTCTACTTTTACTCG 2143
381 .....IleHisAspIleSerIleAsnLeuMetAspIleAsnIle 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2144 GAATCCAAAAATTCACCTACTAATCTCGAATCTAATGATGA...CT 2190
393 uTyrGlyTyrThrAspGluGluIlePheIleValAsnIleGluTyrIle 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2191 CATACACAGAGAGATTTTCACTAGAAAATCTCAAAAACAGTATATAA 2240
410 IeLeuGluIleMetProGlnThrThrIleGlnValAspIleSerGlyIle 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2241 CGAACCAAGAACATG.....CAATTAATAATCTCTCAACAA 2275
427 LysIleValSerIleIleAspPheLeuGlySerGluAsn..... 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2276 GTTTTACAGACTTGAACACTTACAGGGCTCTTAAATATATATATAC 2325
440 IysArgThrGlyAlaLeuGlyIleAsnLeuIleAsnIleTyrSerIle 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2326 TAATATGTCGAGATATATAAGCGGTCAACCGAAGAACATACAGAGAG 2375
456 yAspAsnAsnGluLeuSerHisPheAlaThr..... 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2376 TAATATCCCATTAAGCGAAGACAGATTCATCTCAATCGTATATTAAC 2425
467 ...ThrCysSerAspIleSerArgProIleAsnAspLeuValSerGln 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2426 ATTACTTACTTACTTATGATGATTCAGATGACAGATGAGTGGTGTAT 2475
482 sThrThrGlnLeu.....SerAspIleThrSerArgPheAsn 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2476 TTGAGCTTACTTGAATATATTGAGAGTATCTACTTCTCATATCGT. 2524
495 eValIleGluAlaLeuAsnArgPheIleGlnIleTyrAspSerValMet 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2525 .....AAGATTGAAATGAATCTACCAATCATCATCAATATTCTA 2566
  
```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

(1996) Cancer Research 56:5380-5383."

Tel: 301 838 0200
Fax: 301 838 0208


```

81 1leaspthrglu.....leuleuValglyThLeuthrLeuglygltty 95
      |||
583 ACCCTGACCGAATCGCAGCTGAGCCAGCCAGCTGTAAGTACGCGGTTC 632
      |||
95 1lysthrnglyThrThSerThrSerValasnPhetThrAspAlaAlaGlyA 112
      |||
633 TGGTACTGACCATACGCAAGCATTTGTGAAATTGACGCCACTGGAC 662
      |||
112 sPProMetTyrlLeuthrPhe 118
      |||
683 ATCGCTCGCATTTGTTCTTC 702
      |||
seq_name: gb_est14:AA941922

seq_documentation_block:
LOCUS   AA941922      713 bp      mRNA      EST      25-NOV-1998
DEFINITION  LD27265.5prime LD Drosophila melanogaster embryo pot2 Drosophila
melanogaster cDNA clone LD27265 5prime, mRNA sequence.
ACCESSION  AA941922
VERSION    AA941922.1  GI:3102473
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 713)
            Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
            Brokstein,P., Lewis,S. and Rubin,G.M.
            BGGP/HMT Drosophila EST Project
            Unpublished (1997)
JOURNAL   On Jan 19, 1998 this sequence version replaced gi:2284650.
COMMENT   Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 USA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 272 row: F column: 5
            High quality sequence stop: 606.

FEATURES             Location/Qualifiers
     source            1..713
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="LD27265"
                     /clone_1lb="LD Drosophila melanogaster embryo pot2"
                     /sex="male and female"
                     /dev_stage="0 to 24 hours mixed stage embryonic"
                     /lab_host="X11 Blue"
                     /note="Organ: embryo; Vector: pot2; Site_1: EcorI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pot2."

BASE COUNT      167 a      215 c      194 g      137 t
ORIGIN

```

```

353 CGGACTGGACACACACACCTTAAGTTCGCCAGCCAGCCACTGACGAA 402
26 1ylserSerVal1leAla1leAlaLeupheglyThr1leAlaThrala 42
   |||
403 AGCGGAGGCGATGCTGCTCTGCGTCCAGCTGACCTTGCCTATAGC 452
   |||
43 AsnAla1leAspLeuthrAlaSerThrThralaThralaThrLeuValc1 59
   |||
453 ATGGGTGCC.....CTGAGTCCGACTCGCTTAACGAGACCACTTCAG 496
59 1pRoAlaArg1leThrLeuthrThryllysgluGlyAla.....P 72
   |||
497 TCCAGCCACATTC.....TACAGCCCAACTCGCAAGAAACGCA 537
   |||
72 1ollerThrlleMet.....AspAsnGlyAsn 80
   |||
538 AGATCAACAGGATGTCATGAGATTCAGATTCAGCAAGATTCATCAGAC 587
   |||
81 1leaspthrglu.....leuleuValglyThrLeuthrLeuglygltty 95
   |||
588 ACCCTGACCGAATCGCAGCTGAGCCAGCCAGCTGTAAGTACGCGGTTC 637
   |||
95 1lysthrnglyThrThSerThrSerValasnPhetThrAspAlaAlaGlyA 112
   |||
638 TGGTACTGACCATACGCAAGCATTTGTGAAATTGACGCCACTGGAC 687
   |||
112 sPProMetTyrlLeuthrPhe 118
   |||
688 ATCGCTCGCATTTGTTCTTC 707
   |||
seq_name: gb_est14:AA942159

seq_documentation_block:
LOCUS   AA942159      727 bp      mRNA      EST      25-NOV-1998
DEFINITION  LD26225.5prime LD Drosophila melanogaster embryo pot2 Drosophila
melanogaster cDNA clone LD26225 5prime, mRNA sequence.
ACCESSION  AA942159
VERSION    AA942159.1  GI:3101782
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 727)
            Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
            Brokstein,P., Lewis,S. and Rubin,G.M.
            BGGP/HMT Drosophila EST Project
            Unpublished (1997)
JOURNAL   On Jan 19, 1998 this sequence version replaced gi:2284870.
COMMENT   Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 USA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 262 row: C column: 1
            High quality sequence stop: 698.

FEATURES             Location/Qualifiers
     source            1..727
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="LD26225"
                     /clone_1lb="LD Drosophila melanogaster embryo pot2"
                     /sex="male and female"
                     /dev_stage="0 to 24 hours mixed stage embryonic"
                     /lab_host="X11 Blue"
                     /note="Organ: embryo; Vector: pot2; Site_1: EcorI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pot2."

BASE COUNT      173 a      224 c      193 g      137 t
ORIGIN

```



```

21 AsplysHsmethylserleaserValIleAlaIleAlaLeuPheG1 37
231 .....GCGACGCTCATACCAACACCAACACCAAC 259
37 YTHrleAlaThrAlaAlaAlaAspLeuThrleaserThrAlaT 54
||||| ..... ||||| .....
260 CACCATACCC.....ACGCCATTAACCAACCGGTA 288
54 hrAlaThrLeuValGluProAlaArgIleThrLeuThrTyrGluGly 70
||||| ..... ||| .....
289 CGGGACGCTATTCACACCGCCACCGCTACCGTCACT..... 326
71 AlaprolleThrIleMetaspasnGlysnIleAspThrGluLeuVal 87
326 .....
87 IGIYThrleuThrleuGlyGlyTyrLeuThrGlyThrleuThrSerV 104
||| ..... ||| .....
327 ....ACGACCAACATA.....TACACCGCCACCAAGATGATCTCT 363
104 aAlaAspThrAspAlaAlaGlyAspPrometTyr...LeuThrPheThr 119
||||| ..... ||| .....
364 CCTCTCTCCGACGCGCGCGCGCCACCGCGGAGTGTGATCTTCACT 413
120 SerGlnAsp..... 122
414 GACAGTATATTAAGCGCGCTCAAGTCTGTCATTCATCGCTCGATCG 463
123 .....GlyAsnAsnHisGlnPheThrThrLysV 132
||||| ..... ||| .....
464 CATCTTCATGACTGCTCCGAGCGAGGATCAATCAATCAAAACCTTG 513
132 aIleGlyAspSerArgAspPheAspIleSerProLysValasnGly 148
||||| ..... ||| .....
514 TA.....TCTGCCAGTCTCACGATCGTT..... 536
149 GluAsnLeuValIleGlyAspAspVal...ValleuAlaThrGlySerGlnAs 164
||||| ..... ||| .....
537 GATATCTCTCTTCACGATATCGCTACTGAGAGAAGGCTCTCTCGA 586
164 pPhePheValArgSerIleGlySerLysGlyGlyLysLeuAlaGly 181
||||| ..... ||| .....
587 GGGCATGCTCTCACT.....GCCCAACG 612
181 ySTYrThAspAlaValThrVal 188
||| ..... ||| .....
613 AGCCCAACCGCAAAAGCAACGCTA 635

seq_name: gb_est28:AU003803

seq_documentation_block:
LOCUS AU003803 753 bp mRNA EST 19-JAN-1999
DEFINITION AU003803 Bombyx mori p50(Dalzo) Bombyx mori cDNA clone ws00597,
mRNA sequence.
ACCESSION AU003803
VERSION AU003803.1 GI:4161174
KEYWORDS EST.
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Bombycoidea; Bombycidae; Bombyx.
1 (bases 1 to 753)
Mita,K., Morimoto,M., Shinada,T., Okano,K. and Maeda,S.
Establishment of cDNA database of Bombyx mori
Unpublished (1999)
On May 11, 1999 this sequence version replaced gi:477651.
Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmita@nirs.go.jp
PROJECT = 'CREST project by JST'.

```

```

FEATURES
    source          1..753
                    /organism="Bombyx mori"
                    /strain="p50(Dalzo)"
                    /db_xref="taxon:7091"
                    /clone="ws00597"
                    /clone_11b="Bombyx mori p50(Dalzo)"
BASE COUNT      243 a      177 c      220 g      113 t
ORIGIN
alignment_scores:
    Quality: 108.50      Length: 311
    Ratio: 0.719      Gaps: 11
    Percent Similarity: 48.553      Percent Identity: 18.650
alignment_block:
US-08-699-716a-2 x AU003803
Align seg 1/1 to: AU003803 from: 1 to: 753
215 ArgValGluGlnLeuThrGlyHisGlySerSerValLeuGluGluLeuVal 231
||||| ..... ||| .....
6 CGACCCGATCAAGCTTTAAAGAGAGCCAGAGCTGTATGACGAAGCTTT 55
231 IGIleuValLysAspLysAsnIleAspIle.....SerIleYSTYrA 246
||||| ..... ||| .....
56 AGTCTGTACGCGGAGGTCAACACCAACGCTGCTGCCGATCAAGCTCA 105
246 sPROAlaGlyAspSerGluValPheAlaAsnArgValIleThrAspAsp 262
||||| ..... ||| .....
106 ACAAGCTTCACGAAGACTCAATGAGATGATGAATGAACATGATGAGAAA 155
263 ILeuLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThrI 279
||||| ..... ||| .....
156 TCGGCAGAGCTGACCGCTTCATGACTGAG.....AACGAGACACCT 199
279 eLeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyTleLysArgV 296
||||| ..... ||| .....
200 T.....CACAAATTAAAGAGCGCATTAAGCGCG 228
296 aLysGluPheLeuGluSerSerProAsnThrGlnTProGluLeuArgAla 312
||||| ..... ||| .....
229 GTCGGAGCTGCTGGAACAGAGTCACGACGAGGAGAACTCAACGAT 278
313 PheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
||||| ..... ||| .....
279 CTCTCGCTAAGCTG.....GATGA 298
329 pIleuLysValIleValAspSerMetAsnHisHisGlyAspAlaArgS 346
||||| ..... ||| .....
299 GCTG.....CAGCGCAACGCTTAAA 318
346 eLysLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyr 362
||||| ..... ||| .....
319 ACGATGTGGAAGTCACGAAGCCAGCTTGACAGCGCGCAATGAAATTTC 368
363 SerValIleGlnAlaGluIleAsnLysHisLeuSerSerSerGlyThrI 379
||||| ..... ||| .....
369 AAGACTTTAAA...GAGTTCAGATCAAGTACCAAGTCG..... 407
379 eAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGlyT 396
||||| ..... ||| .....
408 .....CACAGCAACGATC..... 422
396 YTHrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuGlu 412
||||| ..... ||| .....
422 ..... 422
413 LysMetProGlnThrThrIleGlnValAspGlySerGluLysIleVal 429
||||| ..... ||| .....
423 .....CAAGCAAGATGAGAGTCCGCGCGCTCCAAAGAGATGAC 463

```

```

429 1ser1lelytaspPheUgLyserLusnlnyArGThGlYAlAleuc 446
      :::::|||||:::::
464 GGGGCCCGAGAGCATC..... 482
446 1yAsnLeuysAsnSerTySerTyArSnLnysApAnsnGlnLeuSer 462
      |||::::::::::AACACATCCAC 494
483 ..... 494
463 HisPhAlThrThrcysSerAspLysSerArgProLeuAnspleuA 479
      ::::::::::|||:::::
495 GAGGACCTGTGC GCCCGCACGAATCAAGAAGAAAGCGTCGCTCCAAGCA 544
479 1serGlnySthrThcInLeuSerSplerThrSerArgPheAnsSera 496
      :::::|||:::::
545 CCAGCAAGCGCAAAGAAATATGCCAC.....AAG 576
496 1al1eGluAlaleuAnArGPhe1lEGLnlstyrASpservAlmetGln 512
      || ||||| ::::: ||| :::::
577 CTTCGTAGGCCCCGCCACGAATCAAGAAGAAAGCGTCGCTCCAAGCA 626
513 ArgLeu.....LeuAspAspThrSerGlyLys 521
      ::||| ||| ||| |||
627 AAAGCTGTGACTTGAGAGCAGATCAGACAAG 659

seq_name: gb_est4:AA263871

seq_documentation_block:
LOCUS       AA263871          669 bp            mRNA                    EST           27-NOV-1998
DEFINITION   LD07129.5prIme LD Drosophila melanogaster embryo Bluescript
              Drosophila melanogaster cDNA clone LD07129 5prime, mRNA sequence.
ACCESSION    AA263871
VERSION      AA263871.1 GI:2789712
SOURCE       fruit fly.
ORGANISM     Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
              Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
              Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1 (bases 1 to 669)
AUTHORS      Harvey,D., Hong,L., Evans,Holl,M., Pendleton,J., Su,C.,
              Broksrein,P., Lewis,S. and Rubin,G.M.
TITLE        BOSP/HHMT drosophila EST Project
JOURNAL      Unpublished (1997)
COMMENT      On Jan 17, 1998 this sequence version replaced gi:1899948.
              Contact: Harvey, D.
              G. M. Rubin-Molecular and Cell Biology
              University of California Berkeley
              538 LSA, Berkeley, CA 94720-3200, USA
              Fax: 510 643 9947
              Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
              Plate: 71 row: C column: 5
              High quality sequence stop: 539.

FEATURES             source
     location/qualifiers
         1..669
             /organism="Drosophila melanogaster"
             /db_xref="BDBP_EST:BDcln006520"
             /db_xref="taxon:7227"
             /clone="LD07129"
             /clone_lib="LD Drosophila melanogaster embryo Bluescript"
             /sex="male and female"
             /dev_stage="0 to 24 hours mixed stage embryonic"
             /lab_host="SOLR"
             /note="Organ: embryo; Vector: Bluescript SK; Site: 1;
                     EcoRI, Site_2: XhoI; Constructed using Stratagene ZAP-cDNA
                     Synthesis kit. Oligo dt-primed and directionally cloned at
                     EcoRI and XhoI in Bluescript SK(+/-)"

BASE COUNT      158 a      204 c      183 g      124 t
ORIGIN
alignment_scores:
Quality: 107.50      Length: 137
Ratio: 1.581          Gaps: 7

```

```

Percent Similarity: 49.635      Percent Identity: 27.737

alignment_block:
US-08-699-716A-2 x AA263871 ..

Align seg 1/1 to: AA263871 from: 1 to: 669

2 GYHSHSHSHSHSHSHSHSHSHSHSHSHSHSHSHSesSerSeGclYHSHleas 18
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
304 GGACATCATGGCCACACACCATTAAGTGGCCAGCAGCGCTGACGADA 403
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 PAsPaSPAsPlySHsMet.....LYsL 26
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
354 CGGACCTGGACACACACACCTTAAGTGGCCAGCAGCGCTGACGADA 403
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
26 YsLSeSeSeYalIleAlaIleAlaLeuPheGlyThrIleAlaThraLa 42
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
404 AGCGCGAGCGATGCTGCTGCGCTGCTGCACATGACCTTGCTATAGCC 453
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43 ASnAlAlAsPlLeuThraLaseThrThraIAlaThraLThleuVal 59
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
454 ATGGGGGCC.....CTGGAGTCGCGACTGCTTAAGAGACACACTTC 497
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
59 UPtAlAlAgIleThrLeuThrTyrlYsgLugIyAla.....P 72
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498 TCAGGCCACATTC.....TACAAGCCCAACTGCCAGAGAAACGA 538
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 rOlIeThrIleMetAsPaSngIyAsnIleAsPThrGluLeuValGly 88
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
539 AGATCACACGAGATG.....TCATGAGATGCAATTC..... 570
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
89 ThrLeuThrLeuGlyGlyTyrlYsThrGlyThrThrSerThsValAs 105
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
571 .....ACCAAGATATACATACACAG 590
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 nPhetrrAsPaAlaAlaGlyAsPProMetTyrlLeuThrPheThr...SeG 121
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
591 CTTACACGATCGCCACTGACGCCAGCCTGCTTAAGTACAGCGGTTCG 640
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 InAsPglYAsn 124
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
641 CTACTGTGAAC 651

seq_name: gb_est14:AA979134

seq_documentation_block:
LOCUS AA979134 708 bp mRNA EST 24-NOV-1998
DEFINITION LB33408.Sprime LD Drosophila melanogaster embryo P0T2 Drosophila
melanogaster cDNA clone LD33408 5prime, mRNA sequence.
ACCESSION AA979134
VERSION AA979134.1 GI:3156539
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Neoptera; Diptera; Endopterygota; Diptera; Brachycera;
Muscophora; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 708)
Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Brookstein,P., Lewis,S. and Rubin,G.M.
BDGP/HMT Drosophila EST Project
Unpublished (1997)
On Mar 10, 1998 this sequence version replaced gi:2948339.
Contact: Harvey, D.
G.M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
Plate: 334 row: A column: 846.
High quality sequence stop: 546.
Location/Qualifiers
1..708

```



```

|||||
1 CACCATCATACCAACCAACCAAGCTCAGCGTCCCGATGCTACCA 50
8 .....HISHISHISHISerSerglyHisI 17
51 TAGCAGCAGCATGGGACCGCTCATCGCTCATATGATGCA 100
17 leasp..AspAspAspLysHisMetLysLys..IleSerValIleAl 32
101 GGGACCATCTCGGGCATGCGCCACTTAAGATGTCATGACACATCAT 150
32 allelaLeupheglyThrIleAlaThrAlaAsnAlaAlaSerLeuTha 49
151 AGCCGGGAACATGGACACCATTA.....GCTCATCTCAGGACATGGCCA 194
49 laSerThrThr.....AlaThrAlaThrLeuValGlu 59
195 GAGGACCATCTCTCTCTCCGAGCGGTACACACAGCAGCATGTC... 241
60 ProAlaArgIleThrLeuThrThrLysGluGlyAlaProIleThr... 11 75
242 ACCAACAAGATCACCAGGTC.....ATCACATGAT 273
75 emetAspAsnGlyAsnIleAspThrGluLeuValGlyThrLeuThrL 92
274 CATGGACAC.....ATCACGCTCGCATGGCTCCACCGACCA 314
92 euGlyGlyLys.....ThrGlyThrThrSerThr 102
315 CTCACGCTCTCGGAGCGCATCATCAGCAGTACGGGACCAAGACAC 364
103 Ser 103
365 ACA 367

seq_name: gb_est7:AA439056

seq_documentation_block:
LOCUS AA439056 499 bp mRNA EST 28-NOV-1998
DEFINITION LD13501.5prime LD Drosophila melanogaster embryo Bluescript
Drosophila melanogaster cDNA clone LD13501 5prime, mRNA sequence.
ACCESSION AA439056
VERSION AA439056.1 GI:2150937
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 499)
Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Brokstein,P., Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (1997)
Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 USA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 135 row: A column: 1
High quality sequence stop: 427.
Location/Qualifiers
1..499
/organism="Drosophila melanogaster"
/db_xref="BDGP_EST:BDcIn012732"
/db_xref="taxon:7227"
/clone="LD13501"
/clone_lib="LD Drosophila melanogaster embryo Bluescript"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"
/note="Organ: embryo; Vector: Bluescript SK; Site_1:"

```

```

BASE COUNT 107 a 154 c 145 g 93 t
ORIGIN
alignment_scores:
Quality: 105.50 Length: 82
Ratio: 2.638 Gaps: 3
Percent Similarity: 48.780 Percent Identity: 34.146
alignment_block:
US-08-699-716A-2 x AA439056 ..
Align seg 1/1 to: AA439056 from: 1 to: 499
2 GlyHISHISHISHISHISHISHISHISerSerglyHisIleAs 18
|||||
300 GGGCATCATGGCCACACACCATCACCACACCAAGTCCACAGTGA 349
18 pAspAspLysHisMetLysLysIleSerValIleAlaIleAla 35
350 CGGACTGGACACACACCACTTAAGTGGCA..... 380.
35 eupheGlyThrIleAlaThrAlaAsnAlaAlaSerLeuThrAlaSerThr 51
381 .....GCCAGCATCTGAGGAAAGCG 404
52 ThrAlaThrAlaThrLeuValGluProAlaArgIleThrLeuThrL 68
|||||
405 GAGCGATGCTGCTCTGCTGCTGCA.....CTGACCTTGCTATACG 448
449 CATGGGTGCC.....CTGGATTCGACTGCTTAACGAGGACAC 488

seq_name: gb_est28:AJ388692

seq_documentation_block:
LOCUS AJ388692 503 bp mRNA EST 21-MAR-2000
DEFINITION AJ388692 Medicago truncatula R108 Medicago truncatula cDNA clone
M1N0051 similar to glycine-rich RNA binding protein, mRNA sequence.
ACCESSION AJ388692
VERSION AJ388692.1 GI:6603958
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales;
Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 503)
Goyrey,J., Vaubert,D., Jimenez-Zurdo,J.I., Charon,C.,
Trousard,L., Kondorosi,A. and Kondorosi,E.
Analysis of Medicago truncatula nodule expressed sequence tags
Mol Plant Microbe Interact. 13 (1), 62-71 (2000)
20120379
On Apr 3, 1998 this sequence version replaced gi:3018972.
Contact: Goyrey J
Institut des Sciences Vegetales
Centre National de la Recherche Scientifique (CNRS)
Av. de la Terrasse Bat.23, Gif-sur-Yvette, 91198, FRANCE
Additional information about these EST clones can be downloaded
from http://www.cnrs-gif.fr/lsv/AK/index.html. (URL provided by the
author).
Location/Qualifiers
1..503
/organism="Medicago truncatula"
/strain="R108"
/db_xref="taxon:3880"
/clone="M1N0051"
/clone_lib="Medicago truncatula R108"
/tissue_type="symbiotic root nodule"

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 14:58:47 ; Search time 1417.04 Seconds
(without alignments)
1972.304 Million cell updates/sec

Title: US-08-699-716a-1
Perfect score: 1566
Sequence: 1 ATGGGCATCATCATCATCA.....ATGCACAGTCTGTAAATGA 1566

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 972840 segs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_cm:*
- 4: gb_ov:*
- 5: gb_pat:*
- 6: gb_ph:*
- 7: gb_pl1:*
- 8: gb_pl2:*
- 9: gb_pl:*
- 10: gb_pr2:*
- 11: gb_pr3:*
- 12: gb_ro:*
- 13: gb_sts:*
- 14: gb_sy:*
- 15: gb_un:*
- 16: em_fun:*
- 17: em_hum1:*
- 18: em_hum2:*
- 19: em_in:*
- 20: em_cm:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_sy:*
- 29: em_un:*
- 30: em_v1:*
- 31: gb_htg1:*
- 32: gb_htg2:*
- 33: gb_hum4:*
- 34: gb_in1:*
- 35: gb_in2:*
- 36: em_ba1:*
- 37: em_ba2:*
- 38: em_hum3:*
- 39: gb_pr4:*
- 40: gb_htg3:*
- 41: gb_htg4:*
- 42: gb_htg5:*
- 43: gb_htg6:*

44: gb_htg7:*

45: em_htg1:*

46: em_htg2:*

47: em_htg3:*

48: em_hum5:*

49: gb_pl3:*

50: gb_pr5:*

51: gb_htg8:*

52: gb_htg9:*

53: gb_htg10:*

54: gb_htg11:*

55: gb_htg12:*

56: gb_htg13:*

57: gb_htg14:*

58: gb_in3:*

59: gb_htg15:*

60: gb_htg16:*

61: gb_htg17:*

62: em_htg4:*

63: em_htg5:*

64: em_htg6:*

65: em_htg7:*

66: em_hum6:*

67: gb_htg18:*

68: gb_htg19:*

69: gb_htg20:*

70: gb_htg21:*

71: gb_htg22:*

72: gb_htg23:*

73: gb_htg24:*

74: gb_htg25:*

75: gb_htg26:*

76: gb_htg27:*

77: gb_htg28:*

78: gb_htg29:*

79: gb_htg30:*

80: gb_htg31:*

81: gb_v11:*

82: gb_v12:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	1473.6	94.1	1530	5	A56814	A56814 Sequence 22
2	1410.6	90.1	1462	5	A56808	A56808 Sequence 16
3	979.4	62.5	2100	1	YEPICR	M26405 Yersinia pe
4	979.4	62.5	70305	2	YPCDI	AL117189 Yersinia
5	979.4	62.5	70504	2	AF053946	AF053946 Yersinia
6	979.4	62.5	70559	2	AF074612	AF074612 Yersinia
7	976.8	62.4	1014	5	A46411	A46411 Sequence 1
8	976.8	62.4	1014	5	A56793	A56793 Sequence 1
9	974.8	62.2	1014	5	A46413	A46413 Sequence 3
10	974.8	62.2	1014	5	A56795	A56795 Sequence 3
11	974.6	62.2	981	2	AF167310	AF167310 Yersinia
12	967.4	61.8	975	2	AF167309	AF167309 Yersinia
13	957	61.1	981	1	YEPPIVANT	X96802 Yersinia
14	953.8	60.9	2201	1	YEPICRGVHP	X96802 Y.pseudotub
15	927.4	59.2	69673	2	AF102990	M57893 Yersinia ps
16	916.2	58.5	975	1	YEP66PVANT	AF102990 Yersinia
17	914.6	58.4	975	1	YEP108VANT	X96801 Y. enterocol
18	913	58.3	975	1	YEP527VANT	X96801 Y. enterocol
19	911.4	58.2	975	1	YEP080VANT	X96800 Y. enterocol
20	890.4	56.9	1002	1	YEP314VANT	X96798 Y. enterocol
21	882.4	56.3	1002	1	YENCIVANT	X96797 Y. enterocol
22	516.8	33.0	547	5	A56812	A56812 Sequence 20
23	516.8	33.0	5383	1	YPCAF	X61996 Y. pestis ge
24	516.8	33.0	96210	2	YPPMT1	AL117211 Yersinia

c	25	516.6	33.0	100984	2	AF0539347	AF0539347	Yersinia
c	26	516.8	33.0	100990	2	AF074611	AF074611	Yersinia
c	27	217.6	13.9	240	1	S38727	S38727	lcrGV oper
c	28	75	4.8	4168	2	AF010149	AF010149	Pseudomon
c	29	74	4.7	660	14	AF081364	AF081364	Synthetic
c	30	53.4	3.4	521	10	S75264	S75264	WT1-'Wlms'
c	31	53.4	3.4	1402	5	AR000030	AR000030	Sequence
c	32	48.2	3.1	85779	8	SE011856	SE011856	Saccharom
c	33	48	3.1	7218	5	I66494	I66494	Sequence
c	34	47.8	3.1	252102	60	AC005140	AC005140	Plasmodiu
c	35	47.8	3.1	256172	41	AC005139	AC005139	Plasmodiu
c	36	46.8	3.0	110000	31	PFMAL4P1_0	PFMAL4P1_0	Plasmodiu
c	37	46.8	3.0	293431	31	PFMAL13E4	PFMAL13E4	Plasmodiu
c	38	46.4	3.0	1158	33	AF044866	AF044866	Phoebs s
c	39	45.8	2.9	157340	43	AC007926	AC007926	Typanoso
c	40	45.4	2.9	282806	60	AC006279	AC006279	Plasmodiu
c	41	45.2	2.9	21302	19	PFSC4008	PFSC4008	Plasmodiu
c	42	45.2	2.9	224448	29	PFMAL4P4	PFMAL4P4	Plasmodiu
c	43	44.8	2.9	1978	33	DD1DKA	DD1DKA	D. discoidem
c	44	44.8	2.9	12029	34	AE001426	AE001426	Plasmodiu
c	45	44.4	2.8	321003	31	PFMAL4P3	PFMAL4P3	Plasmodiu

ALIGNMENTS

RESULT	1
LOCUS	A56814
DEFINITION	A56814 1530 bp DNA
ACCESSION	Sequence 22 from Patent WO9628551.
VERSION	A56814
KEYWORDS	A56814.1 GI:3712827
SOURCE	Yersinia pestis.
ORGANISM	Yersinia pestis. Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia. 1 (bases 1 to 1530) Tibball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C., Bennett and Alice,M. VACCINES FOR PLAGUE Patent: WO 9628551-A 22 19-SEP-1996; SECR DEFENCE (GB) Other publication AU 4951196 961002. Other publication AU 4951196 961002. Location/Qualifiers 1. 1530 /organism="Yersinia pestis" /db_xref="taxon:632" 13. 1518

BASE COUNT	500 a	296 c	311 g	423 t
ORIGIN				

Query Match	94.11;	Score 1473.6;	DB 5;	Length 1530;
Best Local Similarity	99.11;	Pred. No. 0;		
Matches 1495;	Conservative 0;	Mismatches 4;	Indels 9;	Gaps

Db	11	ATATGAAAAAATCAGTTCGGTTATGCGCATTCGATTTTGGAACTATGCAACTGGTA	70
QY	128	ATGCGCGAGATTTAATGCAAGCACACTGCAACGGCAACTTGTGACCAAGCCGCA	187
Db	71	ATGGCGAGATTTAACGTGCAGACGCCACTGCAACGGCAACTCTTTGAAACAGCCGCA	130
QY	188	TCACCTTACATATAAGGAAGGGGCTCCAAATTTCATTTATGACAAATGGAAATCGATA	247
Db	131	TCACCTTACATATAAGGAAGGGGCTCCAAATTTCATTTATGACAAATGGAAATCGATA	190
QY	248	CAGATTACTTGTGTGACCTTACTCTTGCGCGCTATATAACAGAACCACTACACAT	307
Db	191	CAGATTACTTGTGTGACCTTACTCTTGCGCGCTATATAACAGAACCACTACACAT	250
QY	308	CTGTAACTTACAGATGCGCGGGGTGATCCACTGTACTTAACTTACTTCTCAGATG	367
Db	251	CTGTAACTTACAGATGCGCGGGGTGATCCACTGTACTTAACTTACTTCTCAGATG	310
QY	368	GAATTAACCAACCAATTCATCTACAAAGATGTGGCAAGATTCGTAGATTTGTATCT	427
Db	311	GAATTAACCAACCAATTCATCTACAAAGATGTGGCAAGATTCGTAGATTTGTATCT	370
QY	428	CTCCTAAGTAAACGGGAGAACCTTGTGGGGATACGCTCTTGCGTACGGCAGCC	487
Db	371	CTCCTAAGTAAACGGGAGAACCTTGTGGGGATACGCTCTTGCGTACGGCAGCC	430
QY	488	AGGATTTCTTGTTCGCTCAATTGGTTCACAAAGCGGTAACCTGCAGCAGGTAATACA	547
Db	431	AGGATTTCTTGTTCGCTCAATTGGTTCACAAAGCGGTAACCTGCAGCAGGTAATACA	490
QY	548	CTGATGCTGTAAACGCTAACCGGTATCTAACCAAGATTTCT-----GATTAAGCT	598
Db	491	CTGATGCTGTAAACGCTAACCGGTATCTAACCAAGATTTCTATAAAGCT	550
QY	599	ACGAACAAACCCCAACATTTATTAGAGATGTACAAAGTTAGGGGGAAACACTTA	658
Db	551	ACGAACAAACCCCAACATTTATTAGAGATGTACAAAGTTAGGGGGAAACACTTA	610
QY	659	CTGGTCATGTTCTTCAGTTTGTAGAAATTTGGTTCAGTTAGTCAAAAGTAAATATAG	718
Db	611	CTGGTCATGTTCTTCAGTTTGTAGAAATTTGGTTCAGTTAGTCAAAAGTAAATATAG	670
QY	719	ATATTTCCATTAAATATGATGATCCAGAAAGATTCGAGGTTTTGGCAATAGAGTAATTA	778
Db	671	ATATTTCCATTAAATATGATGATCCAGAAAGATTCGAGGTTTTGGCAATAGAGTAATTA	730
QY	779	CTGATGATTCGAAATTCGCTCAAGAAATCCAGCTATTTCTACCCGAGATACCATTTC	838
Db	731	CTGATGATTCGAAATTCGCTCAAGAAATCCAGCTATTTCTACCCGAGATACCATTTC	790
QY	839	TTAAGGCGGTCATTATGACAAACCACTGCAGAAATGCGATCAAGGAGTAAAGAGTTCC	898
Db	791	TTAAGGCGGTCATTATGACAAACCACTGCAGAAATGCGATCAAGGAGTAAAGAGTTCC	850
QY	899	TTGATCATGCGCGAATACAAATGCGAATTCGGGCGCTTACATGCGAGTATGATTTCT	958
Db	851	TTGATCATGCGCGAATACAAATGCGAATTCGGGCGCTTACATGCGAGTATGATTTCT	910
QY	959	CTTTAACCGCGATCCGATATGATGATGATATTTTGGAAATGATTTGTATTCATATGATC	1018
Db	911	CTTTAACCGCGATCCGATATGATGATGATATTTTGGAAATGATTTGTATTCATATGATC	970
QY	1019	ATCATGCGGATCCGATATGATGATGATATTTTGGAAATGATTTGTATTCATATGATC	1078
Db	971	ATCATGCGGATCCGATATGATGATGATATTTTGGAAATGATTTGTATTCATATGATC	1033
QY	1079	AGATTATTCAGTTATTCAGCCGAATTAATTAAGCATCTGTCTAGTAGTGCCACATTA	1138
Db	1031	AGATTATTCAGTTATTCAGCCGAATTAATTAAGCATCTGTCTAGTAGTGCCACATTA	1099
QY	1139	ATATCCATGATTAATCCATTAACTCATATGATTAATAATTTATATGGTTATACAGATGAAG	1198

Db 1091 ATATCATGATAATCCATTAATCTCATGATGATTAATTTATATGTTATATACATGAC 1150
 QY 1199 AGATTTTAAAGCCAGGCGACAGTACAAAATTTCTGAGAAAATGCTCAACACCATTC 1258
 Db 1151 AGATTTTAAAGCCAGGCGACAGTACAAAATTTCTGAGAAAATGCTCAACACCATTC 1210
 QY 1259 AGGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1318
 Db 1211 AGGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1270
 QY 1319 AAGAACCGGGGCGTGGGTATCTGAAAACATCTCTATTAATAAATAATATG 1378
 Db 1271 AAGAACCGGGGCGTGGGTATCTGAAAACATCTCTATTAATAAATAATATG 1330
 QY 1379 AATATCTACTTTGCCACACCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1438
 Db 1331 AATATCTACTTTGCCACACCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1390
 QY 1439 GCCAAAAACACACTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1498
 Db 1391 GCCAAAAACACACTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1450
 QY 1499 TGAACCGTTTATCAGAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1558
 Db 1451 TGAACCGTTTATCAGAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1510
 QY 1559 GTAATGA 1566
 Db 1511 GTAATGA 1518

RESULT 2
 LOCUS A56808 1462 bp DNA PAT 03-MAR-1998
 DEFINITION Sequence 16 from Patent WO9628551.
 ACCESSION A56808
 VERSION A56808.1 GI:3712821
 KEYWORDS
 SOURCE Yersinia pestis.
 ORGANISM Yersinia pestis.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.
 1 (bases 1 to 1462)
 Tilhail, R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C.,
 Bennett and Alice, M.
 VACCINES FOR PLAGUE
 Patent: WO 9628551-A 16 19-SEP-1996;
 SECUR DEFENCE (GB)
 Other publication ZA 9602036 960716
 Other publication AU 4951196 961002.
 Location/Qualifiers
 1. 1462
 /organism="Yersinia pestis"
 /db_xref="taxon:632"
 <8. 1450
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA03421.1"
 /db_xref="GI:3712822"
 /translation="ADTASTATATATLVEPARITLYKEGAPITIMDNGIDITELLYG
 TLTIGKKTSTSVFTDADPMTLTSTODGNHOFITKVIIGDSDFDISPV
 NGENLVDPVLTAGSDFVRSIGSGLAGKYDATATVTVNSGSTEGRIAYE
 ONOHFTEDLEKRYBOLTHGSSVLELYOLYKDKRIDISIKYDPKDEEVANRIT
 TDIDIELKTLATFLPADATLKGHTDNOLONGIKRKELESSPNTOWELRAFMAY
 HESLTADRIDDIILKTVLDSMNHGARSRLRELALTELKIVSYIOAEIKHLS
 SGTINIDKINIMDKNLGYTDEIFIKAEYKLEKMQDTIIOVGSSEKIVSIND
 FLSSENRTGALNINKNSYSYNDNNELSHFATTCSDKSPINDLVASQKTTQSDITS
 RFSATFALNRFLIOKYSVMORLLDTSGR"

BASE COUNT 476 a 286 c 300 g 400 t
 ORIGIN

Query Match 90.1%; Score 1410.6; DB 5; Length 1462;
 Best Local Similarity 99.1%; Pred. No. 2.2e-295;
 Matches 1432; Conservative 0; Mismatches 4; Indels 9; Gaps 1;

QY 131 CGGAGATTTTACTGCAAGCACCACCTCAACGCGCACTCTTGTGTAACACCGCCGATCA 190
 Db 6 CGGAGATTTTACTGCAAGCACCACCTCAACGCGCACTCTTGTGTAACACCGCCGATCA 65
 QY 191 CTCTTACATATAGAGAGGCGCTCCATTTACATTTAGCAATAGAAACATCGATACAG 250
 Db 66 CTCTTACATATAGAGAGGCGCTCCATTTACATTTAGCAATAGAAACATCGATACAG 125
 QY 251 AATTACTGTTGGTACCTTACTCTTGGCGCTTAAACAGAAACACATAGCACATCTG 310
 Db 126 AATTACTGTTGGTACCTTACTCTTGGCGCTTAAACAGAAACACATAGCACATCTG 185
 QY 311 TTACTTTACAGATCCGCGGGGTATCCCATGTACTTACATTTACTTCTCAGATGGA 370
 Db 186 TTACTTTACAGATCCGCGGGGTATCCCATGTACTTACATTTACTTCTCAGATGGA 245
 QY 371 ATAACACCAATTCAGTACAAAGATGATGGAAGATTTAGAGATTTGATATCTCTC 430
 Db 246 ATAACACCAATTCAGTACAAAGATGATGGAAGATTTAGAGATTTGATATCTCTC 305
 QY 431 CTAAGTAAACGCTGAGAACCTTGTGGGGATGACGTCTGCTTGGCTACGGGACAGCAG 490
 Db 306 CTAAGTAAACGCTGAGAACCTTGTGGGGATGACGTCTGCTTGGCTACGGGACAGCAG 365
 QY 491 ATTCTTTGCTGCTCAATGTTGTTCCAAAGCGGTAAACCTTGACAGAGTAATACACTG 550
 Db 366 ATTCTTTGCTGCTCAATGTTGTTCCAAAGCGGTAAACCTTGACAGAGTAATACACTG 425
 QY 551 ATGCTGTAACGTAACCGTATCTACCAAGATTCAT-----GATTAGACCTACG 601
 Db 426 ATGCTGTAACGTAACCGTATCTACCAAGATTCATGCAAGTGTGATTAGACCTACG 485
 QY 602 AACAAAACCCAAACATTTTATGAGATCTAGAAAAGTTAGGGTGGAAACATCTACTG 661
 Db 486 AACAAAACCCAAACATTTTATGAGATCTAGAAAAGTTAGGGTGGAAACATCTACTG 545
 QY 662 GTCATGTTCTTCACTTTAGAGAAATGCTCAGTTAGTCAAAAGATTAATAATAGATA 721
 Db 546 GTCATGTTCTTCACTTTAGAGAAATGCTCAGTTAGTCAAAAGATTAATAATAGATA 605
 QY 722 TTTCATTAATAATGATCCGAGAAATATTCGAGGTTTGGCAATAGATATTACTG 781
 Db 606 TTTCATTAATAATGATCCGAGAAATATTCGAGGTTTGGCAATAGATATTACTG 665
 QY 782 ATGATATGAAATGCTCAAGAAATCTAGCTTATTTCTACCGGAGATACATTTCTA 841
 Db 666 ATGATATGAAATGCTCAAGAAATCTAGCTTATTTCTACCGGAGATACATTTCTA 725
 QY 842 AAGCGGTCATTTATGACAAACCACTGCAAAATGSCATCAACGAGTAAGAGATCTCTG 901
 Db 726 AAGCGGTCATTTATGACAAACCACTGCAAAATGSCATCAACGAGTAAGAGATCTCTG 785
 QY 902 AATATGCGCGAATACACATGGAATGCGGCGTTCAATGCGAGTAATGATTTCTCT 961
 Db 786 AATATGCGCGAATACACATGGAATGCGGCGTTCAATGCGAGTAATGATTTCTCT 845
 QY 962 TAAACCGCGATCGATGATGATGATGATTTTGAAGATGATTTGATTCATGATGATATC 1021
 Db 846 TAAACCGCGATCGATGATGATGATGATTTTGAAGATGATTTGATTCATGATGATATC 905
 QY 1022 ATGTGTATGCCGTAGCAAGTGGTGGTGAAGATTTAGCTTACCGCGGAATTAAGA 1081
 Db 906 ATGTGTATGCCGTAGCAAGTGGTGGTGAAGATTTAGCTTACCGCGGAATTAAGA 965
 QY 1082 TTTATTAGTTATCAAGCGGAATTAATAAGATCTGTCTAGTAGTGACCATTAATA 1141
 Db 966 TTTATTAGTTATCAAGCGGAATTAATAAGATCTGTCTAGTAGTGACCATTAATA 1025
 QY 1142 TCATGATTAATCATTAATCTCATGATAAAATTTATATGTTATACAGATGAAGA 1201

QY 1186 TATACAGATGAGAGATTGTTAAAGCCAGCCAGAGTACAAATTCGAGAAATGCT 1245
 |||||||
 Db 1081 TATACAGATGAGAGATTGTTAAAGCCAGCCAGAGTACAAATTCGAGAAATGCT 1140
 |||||||
 QY 1246 CAATACACATTCAGAGTGGATGGAGGAGAGAAAAAATAGTCTCGATAAGACTTTCT 1305
 |||||||
 Db 1141 CAATACACATTCAGAGTGGATGGAGGAGAGAAAAAATAGTCTCGATAAGACTTTCT 1200
 |||||||
 QY 1306 GGAATGATGATTAAGAACCGGGCGCTGGTATATCTGAAAAATCTACTCTTATAT 1365
 |||||||
 Db 1201 GGAATGATGATTAAGAACCGGGCGCTGGTATATCTGAAAAATCTACTCTTATAT 1260
 |||||||
 QY 1366 AAGATATATATGATATATCTCATCTTGCACACACTCTCGATTAATCCAGCGCTC 1425
 |||||||
 Db 1261 AAGATATATATGATATATCTCATCTTGCACACACTCTCGATTAATCCAGCGCTC 1320
 |||||||
 QY 1426 AAGACTTGGTTAGCCAAAAAACAACATCAGCTGTCTGATATATCAATCAAGTTTATTC 1485
 |||||||
 Db 1321 AAGACTTGGTTAGCCAAAAAACAACATCAGCTGTCTGATATATCAATCAAGTTTATTC 1380
 |||||||
 QY 1486 GCTATGAGACACGACACCGCTTCTATGAGAAATGATTCAGTATGCAACGCTGCTA 1545
 |||||||
 Db 1381 GCTATGAGACACGACACCGCTTCTATGAGAAATGATTCAGTATGCAACGCTGCTA 1440
 |||||||
 QY 1546 GATGACACGCTGTGTAATGA 1566
 |||||||
 Db 1441 GATGACACGCTGTGTAATGA 1461
 |||||||

RESULT 4

YPCD1 70305 bp DNA BCT 22-MAR-2000
 Yersinia pestis plasmid pCD1.
 AL117189
 AL117189.1 GI:5832423
 chapter: cytotoxic effector; IS100; IS1616; IS1617; lcr;
 low-calcium response; ysc; targeted effector; transposase; type III
 secretion; Y antigen; virulence; ylp; yop; ysc.

SOURCE

ORGANISM

Yersinia pestis
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.

REFERENCE 1 (bases 1 to 70305)

AUTHORS Karlyshev,A.V. and Wren,B.W.

REFERENCE 2 (bases 1 to 70305)

AUTHORS Baker,S.G. and Mungall,K.

REFERENCE 3 (bases 1 to 70305)

AUTHORS James,K.D., Parkhill,J., Barrell,B.G. and Rajandream,M.A.

REFERENCE 4 (bases 1 to 70305)

AUTHORS Submitted (03-SEP-1999) Yersinia pestis sequencing project, Sanger

REFERENCE 5 (bases 1 to 70305)

AUTHORS Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA

REFERENCE 6 (bases 1 to 70305)

AUTHORS E-mail: barrell@sanger.ac.uk DNA supplied by Dr. Andrew Karlyshev

REFERENCE 7 (bases 1 to 70305)

AUTHORS and Prof. Brendan Wren. [3]. Department of Infectious and Tropical

REFERENCE 8 (bases 1 to 70305)

AUTHORS Diseases, London School of Hygiene and Tropical Medicine, Keppel

REFERENCE 9 (bases 1 to 70305)

AUTHORS Street, London WC1E 7HT

REFERENCE 10 (bases 1 to 70305)

AUTHORS Notes:

REFERENCE 11 (bases 1 to 70305)

AUTHORS Yersinia pestis sequencing at The Sanger Centre is funded by

REFERENCE 12 (bases 1 to 70305)

AUTHORS Beowulf Genomics

REFERENCE 13 (bases 1 to 70305)

AUTHORS Details of Y. pestis sequencing at the Sanger Centre are available

REFERENCE 14 (bases 1 to 70305)

AUTHORS on the World Wide Web.

REFERENCE 15 (bases 1 to 70305)

AUTHORS (URL: http://www.sanger.ac.uk/Projects/Y_pestis/)

where these have been used to deduce the initiation codon.
 CAUTION: We may not have predicted the correct initiation codon.
 Where possible we choose an initiation codon (atg, atg, ttg or
 att) which is preceded by an upstream ribosome binding site
 sequence (optimally 5-13bp before the initiation codon). If this
 cannot be identified we choose the most upstream initiation codon.
 Location/Qualifiers

FEATURES

source

repeat_unit

gene

CDS

misc_feature

/organism="Yersinia pestis"
 /plasmid="pCD1"
 /strain="CO-92 Blover Orientalis"
 /db_xref="taxon:632"
 1. 1956
 /note="IS100 element"
 88. 1110
 /gene="YPCD1.01"
 88. 1110
 /gene="YPCD1.01"
 /note="YPCD1.01", probable transposase, len: 340 aa;
 putative insertion sequence IS100, identical to
 corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946,
 AF074612) (340 aa), fasta scores: opt: 2328 z-score:
 3808.9 E(): 0, 100.0% identity in 340 aa overlap. Similar
 to many others e.g. TRA0_ECOLI (EMBL:X14793), lcr, E.coli
 transposase for insertion sequence element IS21 (390 aa)
 (33.1% identity in 329 aa overlap). Contains Pfam match to
 entry PF00239 recombinase, site-specific recombinases,
 score 25.70, E-value 4.8e-06. Contains probable
 helix-turn-helix motif at aa 19-40 (Score 2045, +6.15 SD).
 /codon_start=1
 /transl_table=11
 /label="YPCD1.01"
 /product="putative transposase"
 /protein_id="CAB54878.1"
 /db_xref="GI:5832424"
 /translation="MVFETVMEIKILKMGSSRAIARELGISRNTKRYLQAKSEP
 PKTPRAVAVSLDEEDYRORADHPKIPATVAREIRGSGYGTIRAFR
 SLSPQROEAVAFEPFEGROMOVDMCTMGNSPLIFAVAVIGSYSMYLIETDNR
 YDLETRHRAAFRGVGVREVLNDKNTVYLQDAVOTQHRPHSLPGFKMGFS
 PRICRPFRAQTKGVKERVQYTRNSFTPLMTLRPRGITVDEANRHLRMHDVA
 NQRHETIQARPCDRLMEQSMALPPEKKEYVDHENVNFKRHLHPILSYDS
 FCRGVA"
 112. 195
 /gene="YPCD1.01"
 /note="YPCD1.01", probable transposase, len: 259 aa;
 putative insertion sequence IS100, identical to
 corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946,
 AF074612) (259 aa), fasta scores: opt: 1658 z-score:
 2693.4 E(): 0, 100.0% identity in 259 aa overlap. Similar
 to many others e.g. lcr_ECOLI (EMBL:X14793), lcr, E.coli
 transposase for insertion sequence element IS21 (265 aa)
 (47.4% identity in 249 aa overlap). Contains PF00017
 ATP/GTP-binding site motif A (P-loop).
 /codon_start=1
 /transl_table=11
 /label="YPCD1.02"
 /product="putative ATP-binding protein"
 /protein_id="CAB54879.1"
 /db_xref="GI:5832425"
 /translation="MNEIQRALMALAGOLESLISAPALSOAVDQSWYDFLE
 HILHEKRLARHQRQAMVTRMAFPKATVEEDVDFATAPKQSLQSLSTFEN
 ENIVLSPSVGKTHLAINGAVRAGIVRTTAADLQSLTAQRQRRTTLQR
 GNAPRLLIIDEIGYLPFSEDEKLFQVIAKREKSAMLITNLPFGWDQTFAGDA
 ALTSAMDRLIHSHVQVIGESYRLNKRKAGVIAENPE"
 1434. 1457
 /gene="YPCD1.02"
 /note="P800017 ATP/GTP-binding site motif A (P-loop)."

Query Match	Best Local Similarity	Score	DB	Length
Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	99.9%;	979.4;	DB 2;	70305;
Query 586	ATGATTAGACCTTCAGAACAAACCACACATTTTATTGAGGATCTGAAAAAGTTAGG	645		
Dp 22915	ATGATTAGACCTTCAGAACAAACCACACATTTTATTGAGGATCTGAAAAAGTTAGG	22856		
Qy 646	GTGGACCAACTTACGGTCATGGTCTTCAGTTTAGAAGAAATGGTCACTTGTGTCAA	705		
Dp 22855	GTGGACCAACTTACGGTCATGGTCTTCAGTTTAGAAGAAATGGTCACTTGTGTCAA	22796		
Qy 706	GATAAAAATATAGATATTTTCCATTAAATATGATCCACAAAAGATTCGAGGTTTTGGC	765		
Dp 22795	GATAAAAATATAGATATTTTCCATTAAATATGATCCACAAAAGATTCGAGGTTTTGGC	22736		
Qy 766	AATAGAGTAATTAAGATGATGATGGAATGCTCAAGAAATCCTAAGTTATTTCTACCC	825		
Dp 22735	AATAGAGTAATTAAGATGATGATGGAATGCTCAAGAAATCCTAAGTTATTTCTACCC	22676		
Qy 826	GAGGATACCATTCCTTAAAGGGGTCATATATACAAACCACTGCAAAATGGATCAAGCA	885		
Dp 22675	GAGGATACCATTCCTTAAAGGGGTCATATATACAAACCACTGCAAAATGGATCAAGCA	22616		
Qy 886	GTAAGAGATTCCTTGAATCATCGCCGGAATACAAATGGAAATGCGGCGCTTATGCA	945		
Dp 22615	GTAAGAGATTCCTTGAATCATCGCCGGAATACAAATGGAAATGCGGCGCTTATGCA	22556		
Qy 946	GTAATGATTTCTTTTAAACCGCGGATGATCGATGATATTTTGAAGATGTTGT	1005		
Dp 22555	GTAATGATTTCTTTTAAACCGCGGATGATCGATGATATTTTGAAGATGTTGT	22496		
Qy 1006	GATTCATGATTCATCATGATGATGATGCGCGTACGAAGTGCCTGAAGATTAAGTACGTT	1065		
Dp 22495	GATTCATGATTCATCATGATGATGATGCGCGTACGAAGTGCCTGAAGATTAAGTACGTT	22436		
Qy 1066	ACCGCGGATTAAGATTTATTCAGTTTATTAAGCCGAATTAATTAAGCATCTGCTAGT	1125		
Dp 22435	ACCGCGGATTAAGATTTATTCAGTTTATTAAGCCGAATTAATTAAGCATCTGCTAGT	22376		
Qy 1126	AGTGACACCATTAATTCATGATTAATTCATTAATTCATGATGATTAATTAATGAT	1185		
Dp 22375	AGTGACACCATTAATTCATGATTAATTCATTAATTCATGATGATTAATTAATGAT	22316		
Qy 1186	TATACAGATGAAGATTTTAAAGCCAGCGCAGAGTCAAAATTTCTGAGAAAAATGCTT	1245		
Dp 22315	TATACAGATGAAGATTTTAAAGCCAGCGCAGAGTCAAAATTTCTGAGAAAAATGCTT	22256		
Qy 1246	CAAAACCATTCAGTGGTGGAGGAGCAAAAAAATAGTCTGGATTAAGACTTCTT	1305		
Dp 22255	CAAAACCATTCAGTGGTGGAGGAGCAAAAAAATAGTCTGGATTAAGACTTCTT	22196		
Qy 1306	GGAAGTGAATTAAGAAACCGGGCGTTGGGTATCTGAAAAATCTCATCTTATAT	1365		
Dp 22195	GGAAGTGAATTAAGAAACCGGGCGTTGGGTATCTGAAAAATCTCATCTTATAT	22136		
Qy 1366	AAAGATAATTAATGATTAATTCATTTGACCAACCTGCTGGATTAAGTCCAGGCGGCTC	1425		
Dp 22135	AAAGATAATTAATGATTAATTCATTTGACCAACCTGCTGGATTAAGTCCAGGCGGCTC	22076		

OY	1426	AACGCTGGTATGACCAAAAAAACACATGACGTGCTGATATATACATCAAGCTTTATATCA	1485
OY	22075	AACGCTGGTATGACCAAAAAAACACATGACGTGCTGATATATACATCAAGCTTTATATCA	22016
OY	1486	GCATTTGAAGACATGACACCGTTTCATTTCAGAAATATGATTCAGTATGACACGCTGCTA	1545
OY	22015	GCATTTGAAGACATGACACCGTTTCATTTCAGAAATATGATTCAGTATGACACGCTGCTA	21956
OY	1546	GATGACAGCTCTGCTGAATATGA	1566
Db	21955	GATGACAGCTCTGCTGAATATGA	21935
RESULT	5		
LOCUS	AF053946	AF053946	70504 bp
DEFINITION	Yersinia pestis plasmid pCD1, complete plasmid sequence.		DNA
ACCESSION	AF053946		circular BCT
VERSION	AF053946.1		06-OCT-1998
KEYWORDS	GI:2996222		
SOURCE	Yersinia pestis.		
ORGANISM	Yersinia pestis.		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.		
AUTHORS	Hu, P., Elliott, J., McCready, P., Skowronski, E., Ganes, J., Kobayashi, A., Brubaker, R. R. and Garcia, E.		
TITLE	Structural organization of virulence-associated plasmids of Yersinia pestis		
JOURNAL	J. Bacteriol. 180 (19), 5192-5202 (1998)		
MEDLINE	98422474		
REFERENCE	2 (bases 1 to 70504)		
AUTHORS	Hu, P., Elliott, J., McCready, P., Skowronski, E., Ganes, J., Kobayashi, A., Carrano, A. V., Brubaker, R. and Garcia, E.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-MAR-1998) Biology and Biotechnology, Lawrence Livermore National Lab, L452, 7000 East Ave., Livermore, CA 94550, USA		
FEATURES			
source	location/Qualifiers		
	1..70504		
	/organism="Yersinia pestis"		
	/plasmid="pCD1"		
	/strain="KIM"		
	/db_xref="taxon:632"		
	87..1109		
	/codon_start=1		
	/transl_table=1		
	/product="transposase"		
	/protein_id="AAC62546.1"		
	/db_xref="GI:2996223"		
	/translation="MPTPEYEMEKILHKQMSRAIRREIGISNNTYKRILOAKSEPKYPRPRAVSLDEBYRIQRADAPHTIPRAVIREADGSRGGATILKRLFSLSPQDEPAPVRETEREGROMQVDMGWRNKRSPLHFAVLGISRLYIEFDNMEYDTLETKHRNAFFREGVPRVLYLDNMATVYLQDAVQTGQHRPPLSMQGEKESGPRCRPFAQCTGKGVREAVQYTRNSFYPLMTLRPMDITDVERANSHGRMLADVYNORHETIQARPCDRMLEEQSMALPEPKREYDVHLENLVNFDKHPILHPLSYDSECRBVA"		
	1109..1888		
	/codon_start=1		
	/transl_table=1		
	/product="transposase"		
	/protein_id="AAC62557.1"		
	/db_xref="GI:2996234"		
	/translation="MMPELHORIRWALAGLOLESLSAARPLSOAVDQEWSYDPLEHLHDEKIAARQKQAMATRNRAATPAAVATFEEDITFATGAPQKQLOSLSFIERENIVLPGSGVGKTHLAIMQYEAIVRAIKVFTTAADLLQLSTAOQGRYKTTLOFGVNAFRLILIDEGYLLPESQDEAKLFFGVIAKRREKRSAMITSNLPGQMDQITAGDALTAMADRLILHSHVVOIKGESYRLRQRRAGVIAEANPE"		
	complement(1399..2343)		
	/codon_start=1		
	/transl_table=1		
	/product="transposase"		

gene
CDS
/protein_id="AAC62558.1"
/db_xref="GI:2996245"
/translation="WTEQASERRGCRINGISRLSLHYCPRTADIPVEYVLOKLAQ
IPAFGLMELKQSGLPWVVKVIRYIRLLKLNFRKGRKRLPNHPPLAIPLM
NHCWSPVSDMALDGDGRFRFLFNVEIFVNDG"
complement(2379. .2645)
/gene="lcrS"
complement(2379. .2645)
/gene="lcrS"
/codon_start=1
/transl_table=11
/product="low calcium response locus protein S homolog"
/protein_id="AAC62579.1"
/db_xref="GI:2996256"
/translation="MKKARFTETQILRYLKEVEGGRHVKYDGRNGSGEASYYMKSR
YGGEMSSDIKPKMKREERNRLKQMTASLSIDHEILNDVYAKKL"
complement(3193. .3540)
/gene="yscM"
complement(3193. .3540)
/gene="yscM"
/codon_start=1
/transl_table=11
/product="yop proteins translation protein M homolog"
/protein_id="AAC62586.1"
/db_xref="GI:2996263"
/translation="WKINTQSLNQITGYGHGGQGRLTENPLTENSQIATAER
AFANEVEIHNVTALSRHDICLPLRVSNELKQKGAGVIVTGLRTEQLSLSPAKLIL
LEAMRQDTADG"
complement(3765. .4430)
/gene="yscL"
complement(3765. .4430)
/gene="yscL"
/codon_start=1
/transl_table=11
/product="yop proteins translation protein L homolog"
/protein_id="AAC62596.1"
/db_xref="GI:2996273"
/translation="WSTQYGTGYVMPQFVQIIPSNLSLACGLRLRAEDYQSSLTTER
ELISAKQDARKILADAEQVEQKQRLQWQGMDEARLQATLLHEHQLQCCQYRHHV
EQKMEVVLVLRKLIINDYQVAMTLGVREALALVSNQKQVVRVNDGALAREQDI
AKVHDFEISYLEVTNARLDQGGCILEFVGIIDASIDQIALSAISTITGQKK
VTE"
gene
CDS
complement(4376. .5005)
/gene="yscK"
complement(4376. .5005)
/gene="yscK"
/codon_start=1
/transl_table=11
/product="yop proteins translation protein K homolog"
/protein_id="AAC62605.1"
/db_xref="GI:2996282"
/translation="WMENYISFQIRPCPAVYLLEQLPSMRSILPYLPQWRSAHIL
NAALDESLDQDEEPRGICGLPQROSQELLCRGLVLRHEAIRCTYLAIPQDQ
LITLVNQTLEQITVQHELLIGPMPTHQRFLPTEISRTNIGSLAWLAMEPQPO
AMCKRLSLRLPLAPSEFWVAESORPLAQTLCKLVKQVPTCSHLK"
complement(5005. .5739)
/gene="yscJ"
complement(5005. .5739)
/gene="yscJ"
/codon_start=1
/transl_table=11
/product="yop proteins translation lipoprotein J"
/protein_id="AAC62607.1"
/db_xref="GI:2996284"
/translation="MKYKTSLSLILFLFLGCAVDYTGISQKEGNEMALLNQEG
SADKPDNDGKIKLVEESDVAKIADILKRGYHRESSTLDVFPYDGLISSTEEEL
ARLNAAKQKLSRLSEIDGLVARVHVLEEDNNNGKQKVAASASVFIKHADIDQF
DTYIQAOLVANSIEGLAYDRISIVLVSDVROSHTLPNTSILSIQVSESKGHL
IGLSLILLLPVTNLQYFWLQKK"
complement(5746. .6093)
/gene="yscI"
complement(5746. .6093)
/gene="yscI"
complement(5746. .6093)
/gene="yscI"

```

/codon_start-1
/transl_table-11
/product-"Yop proteins translocation protein I homolog"
/protin_id-"AAC62608.1"
/db_xref-"GI:2996285"
/translacion-"MPRIEIAQADVEIITLLELGAPEPTDQINFPDAMSDETOGL
GSHLEVSIDIQSEFKTVKSDHLTKLAVSDNDMLMOWMSLRITIQEELIATKAG
RMSQNVETLSKGS"
complement(6094. .6591)
/gene-"ysch"
complement(6094. .6591)
/gene-"ysch"
/codon_start-1
/transl_table-11
/product-"Yop proteins translocation protein H homolog"
/protin_id-"AAC62547.1"
/db_xref-"GI:2996224"
/translacion-"MTVTLLKRGSTLSMSSQAVSTLQPYASELKTQLENKLSSEAE
KREVLMOQIYASNPDPHAYLEVATVPKALARRGQGVPAIDPELSEVLAQ
FDSFGKRWELILQVLEGIKPNESQVGLPYSELINKELMILLPNSIYDSLNSHQ
IDMDT"
complement(6588. .6935)
/gene-"yscg"
complement(6588. .6935)
/gene-"yscg"
/codon_start-1
/transl_table-11
/product-"Yop proteins translocation protein G homolog"
/protin_id-"AAC62548.1"
/db_xref-"GI:2996225"
/translacion-"MKIKLVLAETALIGTGHYBEANCIAEMHLKEEPAVOLI
RUSLLNRGDYASALQGNKRAYPDLEPWLALCEYRGLGSALESRLNRLARSQDRI
OTFVNMREQLKT"
complement(6937. .7200)
/gene-"yscf"
complement(6937. .7200)
/gene-"yscf"
/codon_start-1
/transl_table-11
/product-"Yop proteins translocation protein F homolog"
/protin_id-"AAC62549.1"
/db_xref-"GI:2996226"
/translacion-"MSNFSGFTKGTIDADIADVAQTLKKPADANKAVNDSIALKDK
PNPALADQHSINMSVYININSTIVSMKLMOGILQKFP"
complement(7201. .7401)
/gene-"ysce"
complement(7201. .7401)
/gene-"ysce"
/codon_start-1
/transl_table-11
/product-"Yop proteins translocation protein E homolog"
/protin_id-"AAC62550.1"
/db_xref-"GI:2996227"
/translacion-"MTQLEBQLHNETVASTITQLEMALTKLKKDMKRGDAKQYQVW
QRESKALSHAIHIVAGDLK"
complement(7398. .8657)
/gene-"yscd"
complement(7398. .8657)
/gene-"yscd"
/codon_start-1
/transl_table-11
/product-"Yop proteins translocation protein D homolog"
/protin_id-"AAC62551.1"
/db_xref-"GI:2996228"
/translacion-"MSWCBFYQGRHGVNELPHGRVYGSPLQSDIYLSNSETAP
VHLVAVDEGRILTSABRLEDEGPPVPLGTLIRAGSCLEVFLLMTYVAVGQPIPE
TLDVPTQKEPEIDRLPRSRIGISVLISLILTLFLGLGHLWREYNODGQVLEQEV
RLRLATYAAKDDVLTSPKKEGPMWLTGYIODNHARLSLONFLESIGIFERLELRME
ELRQAEFLIQRGLRGIEVSLAPQAGMQLONHARESEIQRKIDSLDLAEEVGLIGLGE
ESKVRIRIAGNQRRKLDALBEQFGLSDFTYVAVKELLELRQGVNDEKLNSPNOQOFER

```

Query Match 62.58; Score 979.4; DB 2; Length 70504;
Best Local Similarity 99.98; Pred. No. 4e-202;

Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 586 ATGATTAGACCTACGACAAACCAACACATTTATTGAGATCTAGAAAAAGTTAG 645
|||||
Db 25311 ATGATTAGACCTACGACAAACCAACACATTTATTGAGATCTAGAAAAAGTTAG 25370
OY 646 GTGACACACTTACTGTCATGTTCTTCAGTTTGAAGAATGGTTCAGTAGTCMAA 705
|||||
Db 25371 GTGACACACTTACTGTCATGTTCTTCAGTTTGAAGAATGGTTCAGTAGTCMAA 25430
OY 706 GATAAAAATATATATNTTCCATTAAATATGATCCAGAAAAATTCGGAGTTTGGC 765
|||||
Db 25431 GATAAAAATATATATNTTCCATTAAATATGATCCAGAAAAATTCGGAGTTTGGC 25490
OY 766 AATAGATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 825
|||||
Db 25491 AATAGATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25550
OY 826 GAGATACCATCTTAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 885
|||||
Db 25551 GAGATACCATCTTAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 25610
OY 886 GTAAAAAGTTCTTGAATCATGATGATGATGATGATGATGATGATGATGATGATG 945
|||||
Db 25611 GTAAAAAGTTCTTGAATCATGATGATGATGATGATGATGATGATGATGATGATG 25670
OY 946 GTATGATTTCTCTTTAACCGCGATGATGATGATGATGATGATGATGATGATGATG 1005
|||||
Db 25671 GTATGATTTCTCTTTAACCGCGATGATGATGATGATGATGATGATGATGATGATG 25730
OY 1006 GATTCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 1065
|||||
Db 25731 GATTCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 25790
OY 1066 ACCGCCAATTAAGATTTATTCAGTATTCAGCCGAATTAAGATCTGCTAGT 1125
|||||
Db 25791 ACCGCCAATTAAGATTTATTCAGTATTCAGCCGAATTAAGATCTGCTAGT 25850
OY 1126 AGTGGCACATTAATATTCATGATTAATCATTAATCTCATGATTAATTAATTA 1185
|||||
Db 25851 AGTGGCACATTAATATTCATGATTAATCATTAATCTCATGATTAATTAATTA 25910
OY 1186 TATACAGATGAAGATTTTAAAGCAGGAGATGATGATGATGATGATGATGATG 1245
|||||
Db 25911 TATACAGATGAAGATTTTAAAGCAGGAGATGATGATGATGATGATGATGATG 25970
OY 1246 CAACACCATTTAGTGGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1305
|||||
Db 25971 CAACACCATTTAGTGGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 26030
OY 1306 GGAAGTGAATTAAGAACACGGGGGCTGGTAACTGTGAATAAAGCTATCTTAA 1365
|||||
Db 26031 GGAAGTGAATTAAGAACACGGGGGCTGGTAACTGTGAATAAAGCTATCTTAA 26090
OY 1366 AAGATATATATATATATCTGATGATGATGATGATGATGATGATGATGATGATG 1425
|||||
Db 26091 AAGATATATATATATATCTGATGATGATGATGATGATGATGATGATGATGATG 26150
OY 1426 AAGACTGTGTTGACCAAAAAACACTGAGCTGTGATATTTACATGAGTTTAAT 1485
|||||
Db 26151 AAGACTGTGTTGACCAAAAAACACTGAGCTGTGATATTTACATGAGTTTAAT 26210
OY 1486 GCTATTGAACACTGAACGTTTCATTGAGAAATATGATGATGATGATGATGATG 1545
|||||
Db 26211 GCTATTGAACACTGAACGTTTCATTGAGAAATATGATGATGATGATGATGATG 26270
OY 1546 GATGACACGCTGTGTAATGA 1566
|||||
Db 26271 GATGACACGCTGTGTAATGA 26291

```

RESULT 6
AF074612

LOCUS	AF074612	70559 bp	DNA	circular	BCT	07-APR-2000
DEFINITION	Yersinia pestis plasmid pCD1, complete plasmid sequence.					
ACCESSION	AF074612	M25810				
VERSION	AF074612.1	GI:3822037				
KEYWORDS	Yersinia pestis.					
SOURCE	Yersinia pestis					
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;					
REFERENCE	1 (bases 43318 to 44840)					
AUTHORS	Leung,K.Y. and Straley,S.C.					
TITLE	The yopM gene of Yersinia pestis encodes a released protein having					
JOURNAL	homology with the human platelet surface protein GPIIb/IIIa					
MEDLINE	J. Bacteriol. 171 (9), 4623-4632 (1989)					
REFERENCE	2 (bases 1 to 70559)					
AUTHORS	Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and					
TITLE	Blattner,F.R.					
JOURNAL	DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1					
MEDLINE	of Yersinia pestis KIM5					
REFERENCE	Infect. Immun. 66 (10), 4611-4623 (1998)					
AUTHORS	3 (bases 43318 to 44840)					
TITLE	Straley,S.C.					
JOURNAL	Direct Submission					
MEDLINE	Submitted (26-APR-1993) Microbiology and Immunology, University of					
REFERENCE	Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA					
AUTHORS	4 (bases 1 to 70559)					
TITLE	Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and					
JOURNAL	Blattner,F.R.					
MEDLINE	Direct Submission					
REFERENCE	Submitted (25-JUN-1998) Microbiology and Immunology, University of					
AUTHORS	Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA					
TITLE	On Apr 7, 2000 this sequence version replaced gi:155539.					
COMMENT	Location/Qualifiers					
FEATURES	1. 70559					
SOURCE	/organism="Yersinia pestis"					
	/plasmid="pCD1"					
	/strain="KIM5"					
	/db_xref="taxon:632"					
gene	57. 368					
	/gene="Y0001"					
CDS	57. 368					
	/gene="Y0001"					
	/note="O103; 43 pct identical (0 gaps) to 100 residues of					
	an approx. 200 aa protein GENEPT: g11537126, of_0198					
	Escherichia coli"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69758.1"					
	/db_xref="GI:3822038"					
	/translation="MHQSRGAASRTSLIMROSGYNYVRLARIRREGIASROP					
	KRRYRGKRVSLASPDLLKROKPEPRNWSGTSITIKVNGWCITLALVIDLISF					
	W"					
gene	665. 1033					
	/gene="nuc"					
	/note="Y0002"					
CDS	665. 1033					
	/gene="nuc"					
	/codon_start=1					
	/transl_table=11					
	/product="endonuclease"					
	/protein_id="AAC69759.1"					
	/db_xref="GI:3822039"					
	/translation="MDTKLQHPITMDYPRVNTKSGKRLRCDFVYIHRDDL					
	VANGIPRTVNSKALHDVYIYDGNKTMGSFNFSQAQVQNSNVLLIMGDFTVQ					
	AYLQYQMSRNKKTDMRSSY"					
gene	1171. 1425					
	/gene="repB"					
	/note="Y0003"					
CDS	1171. 1425					
	/gene="repB"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69760.1"					
	/db_xref="GI:3822040"					
	/translation="MSQRTNATSSSRKRVKGNPLGAEKQKRSVSRKRETHAINV					
	FTQNDKANKLLDLCDSGCTQVTEMIERNIQKKAARTMA"					
	complement(1560..1649)					
	/note="antisense RNA"					
	/product="copA"					
	1667. 1741					
	/gene="repA"					
	/note="Y0004"					
	1667. 1741					
	/gene="repA"					
	/codon_start=1					
	/transl_table=11					
	/product="repA translation protein"					
	/protein_id="AAC69761.1"					
	/db_xref="GI:3822041"					
	/translation="MFRKQYLLRLLLPNCISAGRCD"					
	1734. 2600					
	/gene="repA"					
	/note="Y0005"					
	1734. 2600					
	/gene="repA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69762.1"					
	/db_xref="GI:3822042"					
	/translation="MTNHOALFTHHYRGVKNPNEFTREKKTLPCKRLAKAEGEF					
	TSRFDSHVAARSLSLRHPPLLRARADIALOGCFLYDLARLQISINILNAR					
	EGLAIESKGNLSITRAFRSEGLTYOEGYPOGCGIPDITPTPLFSAG					
	LQSDVAVAAARSREVMNQREKQRLPRLEMLLAKARFVRRERSQTERKHH					
	GLKRRARQDVDRTRDIDALVROLTRFELIAGRFVGMIDVRRKARVERMLSR					
	NNYTRLATGAT"					
	1741. 2935					
	/note="oriR"					
	complement(3427..3645)					
	/gene="Y0006"					
	complement(3427..3645)					
	/gene="Y0006"					
	/note="OriF (f72); 42 pct identical (0 gaps) to 33					
	residues of an approx. 216 aa protein GENEPT: g112055297,					
	Imp2 xenopus laevis"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69763.1"					
	/db_xref="GI:3822043"					
	/translation="MRSPYAGCSYTYVYVQLHNDIHKSAIYHKRLHYLTLQOASS					
	ADPVSMPHRGKAPADIVNVDNWR"					
	4758. 5186					
	/gene="Y0008"					
	4758. 5186					
	/gene="Y0008"					
	/note="OriF (o142); 31 pct identical (1 gap) to 48					
	residues of an approx. 104 aa protein GENEPT: g112149940,					
	ori1 P. syringae"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69764.1"					
	/db_xref="GI:3822044"					
	/translation="MINTFTELPRKIASHFGELKSDDEYGLCELILNDVVIMLARA					
	DEILNRLILGILGIFSGEARSASASQIFCYISIALNKDQPCFAMSEIELLILAFKHL					
	SIDELNVNWSKEINAFDMLSLVSLPAPETAPSYISYISIG"					
	5204. 7402					
	/gene="ypkA"					
	/note="Y0009"					
	5204. 7402					
	/gene="ypkA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69765.1"					
	/db_xref="GI:3822045"					
	/translation="MSTKQYLLRLLLPNCISAGRCD"					
	1734. 2600					
	/gene="repA"					
	/note="Y0005"					
	1734. 2600					
	/gene="repA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69762.1"					
	/db_xref="GI:3822042"					
	/translation="MTNHOALFTHHYRGVKNPNEFTREKKTLPCKRLAKAEGEF					
	TSRFDSHVAARSLSLRHPPLLRARADIALOGCFLYDLARLQISINILNAR					
	EGLAIESKGNLSITRAFRSEGLTYOEGYPOGCGIPDITPTPLFSAG					
	LQSDVAVAAARSREVMNQREKQRLPRLEMLLAKARFVRRERSQTERKHH					
	GLKRRARQDVDRTRDIDALVROLTRFELIAGRFVGMIDVRRKARVERMLSR					
	NNYTRLATGAT"					
	1741. 2935					
	/note="oriR"					
	complement(3427..3645)					
	/gene="Y0006"					
	complement(3427..3645)					
	/gene="Y0006"					
	/note="OriF (f72); 42 pct identical (0 gaps) to 33					
	residues of an approx. 216 aa protein GENEPT: g112055297,					
	Imp2 xenopus laevis"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69763.1"					
	/db_xref="GI:3822043"					
	/translation="MRSPYAGCSYTYVYVQLHNDIHKSAIYHKRLHYLTLQOASS					
	ADPVSMPHRGKAPADIVNVDNWR"					
	4758. 5186					
	/gene="Y0008"					
	4758. 5186					
	/gene="Y0008"					
	/note="OriF (o142); 31 pct identical (1 gap) to 48					
	residues of an approx. 104 aa protein GENEPT: g112149940,					
	ori1 P. syringae"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69764.1"					
	/db_xref="GI:3822044"					
	/translation="MINTFTELPRKIASHFGELKSDDEYGLCELILNDVVIMLARA					
	DEILNRLILGILGIFSGEARSASASQIFCYISIALNKDQPCFAMSEIELLILAFKHL					
	SIDELNVNWSKEINAFDMLSLVSLPAPETAPSYISYISIG"					
	5204. 7402					
	/gene="ypkA"					
	/note="Y0009"					
	5204. 7402					
	/gene="ypkA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69765.1"					
	/db_xref="GI:3822045"					
	/translation="MSTKQYLLRLLLPNCISAGRCD"					
	1734. 2600					
	/gene="repA"					
	/note="Y0005"					
	1734. 2600					
	/gene="repA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69762.1"					
	/db_xref="GI:3822042"					
	/translation="MTNHOALFTHHYRGVKNPNEFTREKKTLPCKRLAKAEGEF					
	TSRFDSHVAARSLSLRHPPLLRARADIALOGCFLYDLARLQISINILNAR					
	EGLAIESKGNLSITRAFRSEGLTYOEGYPOGCGIPDITPTPLFSAG					
	LQSDVAVAAARSREVMNQREKQRLPRLEMLLAKARFVRRERSQTERKHH					
	GLKRRARQDVDRTRDIDALVROLTRFELIAGRFVGMIDVRRKARVERMLSR					
	NNYTRLATGAT"					
	1741. 2935					
	/note="oriR"					
	complement(3427..3645)					
	/gene="Y0006"					
	complement(3427..3645)					
	/gene="Y0006"					
	/note="OriF (f72); 42 pct identical (0 gaps) to 33					
	residues of an approx. 216 aa protein GENEPT: g112055297,					
	Imp2 xenopus laevis"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69763.1"					
	/db_xref="GI:3822043"					
	/translation="MRSPYAGCSYTYVYVQLHNDIHKSAIYHKRLHYLTLQOASS					
	ADPVSMPHRGKAPADIVNVDNWR"					
	4758. 5186					
	/gene="Y0008"					
	4758. 5186					
	/gene="Y0008"					
	/note="OriF (o142); 31 pct identical (1 gap) to 48					
	residues of an approx. 104 aa protein GENEPT: g112149940,					
	ori1 P. syringae"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69764.1"					
	/db_xref="GI:3822044"					
	/translation="MINTFTELPRKIASHFGELKSDDEYGLCELILNDVVIMLARA					
	DEILNRLILGILGIFSGEARSASASQIFCYISIALNKDQPCFAMSEIELLILAFKHL					
	SIDELNVNWSKEINAFDMLSLVSLPAPETAPSYISYISIG"					
	5204. 7402					
	/gene="ypkA"					
	/note="Y0009"					
	5204. 7402					
	/gene="ypkA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69765.1"					
	/db_xref="GI:3822045"					
	/translation="MSTKQYLLRLLLPNCISAGRCD"					
	1734. 2600					
	/gene="repA"					
	/note="Y0005"					
	1734. 2600					
	/gene="repA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69762.1"					
	/db_xref="GI:3822042"					
	/translation="MTNHOALFTHHYRGVKNPNEFTREKKTLPCKRLAKAEGEF					
	TSRFDSHVAARSLSLRHPPLLRARADIALOGCFLYDLARLQISINILNAR					
	EGLAIESKGNLSITRAFRSEGLTYOEGYPOGCGIPDITPTPLFSAG					
	LQSDVAVAAARSREVMNQREKQRLPRLEMLLAKARFVRRERSQTERKHH					
	GLKRRARQDVDRTRDIDALVROLTRFELIAGRFVGMIDVRRKARVERMLSR					
	NNYTRLATGAT"					
	1741. 2935					
	/note="oriR"					
	complement(3427..3645)					
	/gene="Y0006"					
	complement(3427..3645)					
	/gene="Y0006"					
	/note="OriF (f72); 42 pct identical (0 gaps) to 33					
	residues of an approx. 216 aa protein GENEPT: g112055297,					
	Imp2 xenopus laevis"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69763.1"					
	/db_xref="GI:3822043"					
	/translation="MRSPYAGCSYTYVYVQLHNDIHKSAIYHKRLHYLTLQOASS					
	ADPVSMPHRGKAPADIVNVDNWR"					
	4758. 5186					
	/gene="Y0008"					
	4758. 5186					
	/gene="Y0008"					
	/note="OriF (o142); 31 pct identical (1 gap) to 48					
	residues of an approx. 104 aa protein GENEPT: g112149940,					
	ori1 P. syringae"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69764.1"					
	/db_xref="GI:3822044"					
	/translation="MINTFTELPRKIASHFGELKSDDEYGLCELILNDVVIMLARA					
	DEILNRLILGILGIFSGEARSASASQIFCYISIALNKDQPCFAMSEIELLILAFKHL					
	SIDELNVNWSKEINAFDMLSLVSLPAPETAPSYISYISIG"					
	5204. 7402					
	/gene="ypkA"					
	/note="Y0009"					
	5204. 7402					
	/gene="ypkA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69765.1"					
	/db_xref="GI:3822045"					
	/translation="MSTKQYLLRLLLPNCISAGRCD"					
	1734. 2600					
	/gene="repA"					
	/note="Y0005"					
	1734. 2600					
	/gene="repA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69762.1"					
	/db_xref="GI:3822042"					
	/translation="MTNHOALFTHHYRGVKNPNEFTREKKTLPCKRLAKAEGEF					
	TSRFDSHVAARSLSLRHPPLLRARADIALOGCFLYDLARLQISINILNAR					
	EGLAIESKGNLSITRAFRSEGLTYOEGYPOGCGIPDITPTPLFSAG					
	LQSDVAVAAARSREVMNQREKQRLPRLEMLLAKARFVRRERSQTERKHH					
	GLKRRARQDVDRTRDIDALVROLTRFELIAGRFVGMIDVRRKARVERMLSR					
	NNYTRLATGAT"					
	1741. 2935					
	/note="oriR"					
	complement(3427..3645)					
	/gene="Y0006"					
	complement(3427..3645)					
	/gene="Y0006"					
	/note="OriF (f72); 42 pct identical (0 gaps) to 33					
	residues of an approx. 216 aa protein GENEPT: g112055297,					
	Imp2 xenopus laevis"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69763.1"					
	/db_xref="GI:3822043"					
	/translation="MRSPYAGCSYTYVYVQLHNDIHKSAIYHKRLHYLTLQOASS					
	ADPVSMPHRGKAPADIVNVDNWR"					
	4758. 5186					
	/gene="Y0008"					
	4758. 5186					
	/gene="Y0008"					
	/note="OriF (o142); 31 pct identical (1 gap) to 48					
	residues of an approx. 104 aa protein GENEPT: g112149940,					
	ori1 P. syringae"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69764.1"					
	/db_xref="GI:3822044"					
	/translation="MINTFTELPRKIASHFGELKSDDEYGLCELILNDVVIMLARA					
	DEILNRLILGILGIFSGEARSASASQIFCYISIALNKDQPCFAMSEIELLILAFKHL					
	SIDELNVNWSKEINAFDMLSLVSLPAPETAPSYISYISIG"					
	5204. 7402					
	/gene="ypkA"					
	/note="Y0009"					
	5204. 7402					
	/gene="ypkA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69765.1"					
	/db_xref="GI:3822045"					
	/translation="MSTKQYLLRLLLPNCISAGRCD"					
	1734. 2600					
	/gene="repA"					
	/note="Y0005"					
	1734. 2600					
	/gene="repA"					
	/codon_start=1					
	/transl_table=11					
	/product="replication protein"					
	/protein_id="AAC69762.1"					
	/db_xref="GI:3822042"					
	/translation="MTNHOALFTHHYRGVKNPNEFTREKKTLPCKRLAKAEGEF					
	TSRFDSHVAARSLSLRHPPLLRARADIALOGCFLYDLARLQISINILNAR					
	EGLAIESKGNLSITRAFRSEGLTYOEGYPOGCGIPDITPTPLFSAG					
	LQSDVAVAAARSREVMNQREKQRLPRLEMLLAKARFVRRERSQTERKHH					
	GLKRRARQDVDRTRDIDALVROLTRFELIAGRFVGMIDVRRKARVERMLSR					
	NNYTRLATGAT"					
	1741. 2935					
	/note="oriR"					
	complement(3427..3645)					
	/gene="Y0006"					
	complement(3427..3645)					
	/gene="Y0006"					
	/note="OriF (f72); 42 pct identical (0 gaps) to 33					
	residues of an approx. 216 aa protein GENEPT: g112055297,					
	Imp2 xenopus laevis"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69763.1"					
	/db_xref="GI:3822043"					
	/translation="MRSPYAGCSYTYVYVQLHNDIHKSAIYHKRLHYLTLQOASS					
	ADPVSMPHRGKAPADIVNVDNWR"					
	4758. 5186					
	/gene="Y0008"					
	4758. 5186					
	/gene="Y0008"					
	/note="OriF (o142); 31 pct identical (1 gap) to 48					
	residues of an approx. 104 aa protein GENEPT: g112149940,					
	ori1 P. syringae"					
	/codon_start=1					
	/transl_table=11					
	/product="unknown"					
	/protein_id="AAC69764.1"					
	/db_xref="GI:3822044"					
	/translation="MINTFTELPRKIASHFGELKSDDEYGLCELILNDVVIMLARA					
	DEILNRLILGILGIFSGEARSASASQIFCYISIALNKDQPCFAMSEIELLILAFKHL					
	SIDELNVNWSKEINAFD					

[illegible]

Db	38099	AATGAGTAAATTACTGATGATATCGAATTTGGTCGAAGAAAACCTAGCTTATTTTTCACCC	38158
QY	826	GAGGATPACCATTTTAAAGCGCGTATTATGACAACCACTGCAAAATGGCATCAAGCGA	885
Db	38159	GAGATGCCATTTCTTAAAGCGCGTATTATGACAACCACTGCAAAATGGCATCAAGCGA	38218
QY	886	GTAAGAAGATCCTTAAATCATCGCGGAATPACAAATGGGAATTGGGGGCTCATGGGA	945
Db	38219	GTAAGAAGATCCTTAAATCATCGCGGAATPACAAATGGGAATTGGGGGCTCATGGGA	38278
QY	946	GTAATGATTTCTCTTTAAACCGCGCATGATATCATGATGATATTTTGAAGATGATTGT	1005
Db	38279	GTAATGATTTCTCTTTAAACCGCGCATGATATCATGATGATATTTTGAAGATGATTGT	38338
QY	1006	GATTCAATGAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1065
Db	38339	GATTCAATGAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	38399
QY	1066	ACGGCGCAATTAAGATTTATTCATCTATTCATTCATTCATTCATTCATTCATTCATTCAT	1125
Db	38399	ACGGCGCAATTAAGATTTATTCATCTATTCATTCATTCATTCATTCATTCATTCATTCAT	38458
QY	1126	AGTGGCCACATTAATATTCATGATTAATTCATTAATTCATTAATTCATTAATTCATTAAT	1185
Db	38459	AGTGGCCACATTAATATTCATGATTAATTCATTAATTCATTAATTCATTAATTCATTAAT	38518
QY	1186	TATPACAGATGAAGATTTTAAAGCCAGCGCAGATGACAAATTTCTGAGAAATGCT	1245
Db	38519	TATPACAGATGAAGATTTTAAAGCCAGCGCAGATGACAAATTTCTGAGAAATGCT	38578
QY	1246	CAAAACCACTTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1305
Db	38579	CAAAACCACTTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	38638
QY	1306	GGAAGTGAATTAAGAACACCGGCGCTGGGGTATCTGAAAACTCATCTCTTATAT	1365
Db	38639	GGAAGTGAATTAAGAACACCGGCGCTGGGGTATCTGAAAACTCATCTCTTATAT	38698
QY	1366	AAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1425
Db	38699	AAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	38758
QY	1426	AAGGACTTGGTTAGCCAAAAACAACCTAGCTGCTGATATTTACATACGTTTAAATTA	1485
Db	38759	AAGGACTTGGTTAGCCAAAAACAACCTAGCTGCTGATATTTACATACGTTTAAATTA	38818
QY	1486	GCTATTGAAGACATGAACGCTTCAATTCAGAAATATGATACAGTAGCAACGCTGCGTA	1545
Db	38819	GCTATTGAAGACATGAACGCTTCAATTCAGAAATATGATACAGTAGCAACGCTGCGTA	38878
QY	1546	GATGACACGCTCTGTGAATGA 1566	
Db	38879	GATGACACGCTCTGTGAATGA 38899	
RESULT	7		
LOCUS	A46411	1014 bp DNA	PAT 07-MAR-1997
DEFINITION	A46411	Sequence 1 from Patent W0524475.	
ACCESSION	A46411		
VERSION	A46411.1	GI:2300612	
KEYWORDS			
SOURCE	Yersinia pestis.		
ORGANISM	Yersinia pestis.		
REFERENCE	1 (bases 1 to 1014)		
AUTHORS	Tidball, R.W., Williamson, E.D. and Leary, S.E.		
TITLE	VACCINE COMPOSITIONS		
JOURNAL	Patent: WO 9524475-A 1 14-SEP-1995;		
COMMENT	SECUR DEFENCE BRIT (GB)		
FEATURES	Other publication AD 1653995 950925.		
	Location/Qualifiers		

Source

CDS

```

source
1. .1014
/organism="Yersinia pestis"
/db_xref="taxon:632"
1. .990
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA02893.1"
/db_xref="gi:2300613"
/transl_table="MSEIRAYEQNPQHIEDLEKRYVEQLTGSGSVLEELVOLVYKDK
KNIDISIKDPRKDSVEFANVRVIDDLELKKILAYFLPEDAIIKGGHYDQNLONGIKR
KVFLEESPTQOMELRALFAMVHFSFLADIDDDILKVIYDSNMHGGDASRKLREEL
ALELAEKIVSYIOAEINKRILSSGGINIHDKSNIMDKNLYGTDEEFASAEYK
LEKRPOTIOWDSQEKRIYSIKDELGSNKRTGALGNKNSYGNKNKNNELSHRATTC
SDKRPLNDLYSQRTTQDLSDTTSRPNSEALEANRFIOKIDSVMORLDTSGK"
346 a 181 c 201 g 286 t

```

Query Match	63.48	Score 976.8	DB 5	Length 1014
Best Local Similarity	99.88	Pred. No. 1.8e-201		
Matches 978; Conservative	0	Mismatches 2	Indels 0	Gaps 0

QY	587	TGTTAAGGCTACGACCAAAACCCACACATTTTATAGAGATCTAGAAAAAGTTAGG	646
Db	11	TCATTAAGGCTACGACCAAAACCCACACATTTTATAGAGATCTAGAAAAAGTTAGG	70
QY	647	TGGAACAATCTACGGGATCGTGTCTCTCACTTTAGAGAAATGGTTCGTTAGTCAAG	706
Db	71	TGGAACAATCTACGGGATCGTGTCTCTCACTTTAGAGAAATGGTTCGTTAGTCAAG	130
QY	707	ATAAAAATATAGATATTTCCATTAAATATATATCCAGAAAAAGATTCGAGATTTTGCA	766
Db	131	ATAAAAATATAGATATTTCCATTAAATATATATCCAGAAAAAGATTCGAGATTTTGCA	190
QY	767	ATAGATTAATTAATCTGATGATATCGAATTCGTCAAGAAAAATCCTAGCTATTTTCAACCG	826
Db	191	ATAGATTAATTAATCTGATGATATCGAATTCGTCAAGAAAAATCCTAGCTATTTTCAACCG	250
QY	827	AGGATACCATTTCTTAAGGGGGTCATTATACACAACCACTGCAAAATGGCATCAAGCAG	886
Db	251	AGGATGCAATCTTAAAGGGGGTCATTATACACAACCACTGCAAAATGGCATCAAGCAG	310
QY	887	TAAAGAGTTCCTGATTCATTCGCCGAAATACAAATGGGAATGGGGGGTTCATGGCAG	946
Db	311	TAAAGAGTTCCTGATTCATTCGCCGAAATACAAATGGGAATGGGGGGTTCATGGCAG	370
QY	947	TAAATGATTTCTCTTTAACCGCGCATGATTCGATGATGATTTTGAAGAGATTTGTG	1006
Db	371	TAAATGATTTCTCTTTAACCGCGCATGATTCGATGATGATTTTGAAGAGATTTGTG	430
QY	1007	ATTCATGATCATCATGATGGTATGCCCGTAGCAAGTGGGTGAAGATTAAGCTGAGCTTA	1066
Db	431	ATTCATGATCATCATGATGGTATGCCCGTAGCAAGTGGGTGAAGATTAAGCTGAGCTTA	490
QY	1067	CCGGCGAATTAAGATTTATTCAGTATTTTAAGCCGAATTAATAGCATCTGTCTAGTA	1122
Db	491	CCGGCGAATTAAGATTTATTCAGTATTTTAAGCCGAATTAATAGCATCTGTCTAGTA	550
QY	1127	GTGGCACCATTAATATTCATGATTAATCCATTATCTCATGATTAATAATTTATATGTT	610
Db	551	GTGGCACCATTAATATTCATGATTAATCCATTATCTCATGATTAATAATTTATATGTT	670
QY	1187	ATACAGATGAAGAGATTTTAAAGCCAGCCAGCAAGTACAAATATTCGAGAAAAATGCCCT	1246
Db	611	ATACAGATGAAGAGATTTTAAAGCCAGCCAGCAAGTACAAATATTCGAGAAAAATGCCCT	670
QY	1247	AAACCCACCATTCAGGTGATGGAGCGAGAAAAAAATAGTCTCGATTAAGGACTTCTTG	1306
Db	671	AAACCCACCATTCAGGTGATGGAGCGAGAAAAAAATAGTCTCGATTAAGGACTTCTTG	730
QY	1307	GAACTGGAATTAAGAAACCGGGGCGTTGGGTATCTGAAAAATCTATCTTATATA	1366

Db	731	GAATGAGAAATAAAAAGAACCGGGGCGTTGGGTATCTGAAAAACATCACTCTTTAATA	790
OY	1367	AAGATAATATGAATATATCTCACTTTGCCACACCGCTGTGGATAAAGTCAGAGCGCTCA	1428
Db	791	AAGATATATTAAGATTTCTCACTTTGCCACACCGCTGTGGATAAAGTCAGAGCGCTCA	850
OY	1427	ACGACTGTGTAGCCAAAACAAACACAGCTGTCATATTAATCAATCAGCTTTAATTAG	1488
Db	851	ACGACTGTGTAGCCAAAACAAACACAGCTGTCATATTAATCAATCAGCTTTAATTAG	910
OY	1487	CTATTGAAGCAGCTGAACCGTTTCATTCAAGAAATATGATTCAGTATGCAACGCTGTAG	1546
Db	911	CTATTGAAGCAGCTGAACCGTTTCATTCAAGAAATATGATTCAGTATGCAACGCTGTAG	970
OY	1547	ATGACACGCTCGGTTAATGA	1566
Db	971	ATGACACGCTCGGTTAATGA	990

RESULT	8			
LOCUS	A56793			
DEFINITION	A56793	1014 bp	DNA	PAT 03-MAR-1998
ACCESSION	A56793	Sequence 1 from Patent WO9628551.		
VERSION	A56793.1	GI:3712808		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				

Query Match	62.4%	Score 976.8	DB 5	Length 1014
Best Local Similarity	99.8%	Pred. No. 1.8e-201		
Matches 978	Conservative	0	Mismatches 2	Indels 0
			Gaps	0
QY 587	TGATTAGAGCCTCAGACAAAAAACCACAACTTTATTGAGAGCTAGAAAAAGTTAGCG	646		
DB 11	TCATTAGAGCCTCAGACAAAAAACCACAACTTTATTGAGAGCTAGAAAAAGTTAGCG	70		
QY 647	TGGAAACAATTACTGGTCATGGTTCTTCAGTTTGAAGAATTGGTTCACTAGTCAAG	706		
DB 71	TGGAAACAATTACTGGTCATGGTTCTTCAGTTTGAAGAATTGGTTCACTAGTCAAG	130		
QY 707	ATAAAAATATGATATTTCCATTAAATATGATCCAGAAAAAGATTGGAGGTTTTTGCA	766		
DB 131	ATAAAAATATGATATTTCCATTAAATATGATCCAGAAAAAGATTGGAGGTTTTTGCA	190		

QY	587	TGATTAGAGCCTACGACAAACCAACCAACATTTTATTATGAGAGCTCGAAAAAGTTGGG	648
Db	11	TCATTAGAGCCTACGACAAACCAACCAACATTTTATTATGAGAGCTCGAAAAAGTTGGG	70
QY	647	TGGAAACACTACTGCGTATGTTGTTCTTCAGTTTAGAAGATTGGTCACTAGTCAAG	708
Db	71	TGGAAACACTACTGCGTATGTTGTTCTTCAGTTTAGAAGATTGGTCACTAGTCAAG	130
QY	707	ATAAAAATATAGATATTTCCATTAAATATGATCCAGAAAAAGATTCGAGAGTTTTCCTCA	766
Db	131	ATAAAAATATAGATATTTCCATTAAATATGATCCAGAAAAAGATTCGAGAGTTTTCCTCA	180

QY 767 ATAGAGTAATCTAGATGATATGCAATGCTCAAGAAAATCTAGTATTTCTACCG 826
 |||||||
 Db 191 ATAGAGTAATCTAGATGATATGCAATGCTCAAGAAAATCTAGTATTTCTACCG 250
 QY 827 AGGATACCATTTCTTAAGGGCGGTATATGCAACCAACTGCAAAATGGCATCAAGGAG 886
 |||||||
 Db 251 AGGATACCATTTCTTAAGGGCGGTATATGCAACCAACTGCAAAATGGCATCAAGGAG 310
 QY 887 TAAAGAGTTCCTGAATCATCGCCGAATACAAATGGGAATTCGGGCGTTTATGGCAG 946
 |||||||
 Db 311 TAAAGAGTTCCTGAATCATCGCCGAATACAAATGGGAATTCGGGCGTTTATGGCAG 370
 QY 947 TAATGATTTCTCTTTAACCAGCATGATGATGATATTTTGAAGTGTGTG 1006
 |||||||
 Db 371 TAATGATTTCTCTTTAACCAGCATGATGATGATATTTTGAAGTGTGTG 430
 QY 1007 ATTCAATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
 |||||||
 Db 431 ATTCAATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 490
 QY 1067 CCGGCAATTAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126
 |||||||
 Db 491 CCGGCAATTAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 550
 QY 1127 GTGGCACCATAAATATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
 |||||||
 Db 551 GTGGCACCATAAATATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610
 QY 1187 ATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246
 |||||||
 Db 611 ATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
 QY 1247 AAACCAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
 |||||||
 Db 671 AAACCAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
 QY 1307 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
 |||||||
 Db 731 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790
 QY 1367 AAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426
 |||||||
 Db 791 AAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850
 QY 1427 ACGACTGTTAGGCAAAAACAACACTGATGATGATGATGATGATGATGATGATGATGAT 1486
 |||||||
 Db 851 ACGACTGTTAGGCAAAAACAACACTGATGATGATGATGATGATGATGATGATGATGAT 910
 QY 1487 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1546
 |||||||
 Db 911 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970
 QY 1547 ATGACAGCTCTGGTAAATGA 1566
 |||||||
 Db 971 ATGACAGCTCTGGTAAATGA 990
 RESULT 9
 A46413 1014 bp DNA PAT 07-MAR-1997
 LOCUS A46413
 DEFINITION Sequence 3 from Patent WO9524475.
 ACCESSION A46413
 VERSION A46413.1 GI:2300614
 KEYWORDS
 ORGANISM
 SOURCE
 Yersinia pestis.
 Yersinia pestis.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.
 REFERENCES
 1 (bases 1 to 1014)
 Tishball, R.W., Williamson, E.D. and Leary, S.E.
 VACCINE COMPOSITIONS
 JOURNAL Patent: WO 9524475-A 3 14-SEP-1995;

COMMENT SECR DEFENCE BRIT (GB)
 FEATURES Other publication AU 1853995 950925.
 source Location/Qualifiers
 1..1014
 /organism="Yersinia pestis"
 /db_xref="taxon:632"
 1..987
 /note="unnamed protein product; protein sequence is in
 conflict with the conceptual translation"
 /codon_start=1
 /transl_table=11
 /protein_id="CA02894.1"
 /db_xref="GI:2300615"
 /translation="GIPGIRAVEONPOHFIEDLEKRYEOLTHGSSVLELVOLVKD
 KNDISIKYDPRDSEVFANRYTDDIELKRLAVFLPEDALIKGHYDNLONGIK
 RKPELESPNOMELRANMAVMSLSLADRIDDLIKYVDSNMHHGDAKRLREL
 AELRAELKYSYIOAEINKHLSSTGTHHDSINLMDKNLYGYDEIFRASAETKI
 LEMPOTTIYDGESEKIVSINDFLESSEKRRGALGNLSYKNDNELSHFATTC
 SDRSLNDVDSQKTTLSLDRSRENSALEALNRIQKYSVMORLDDTSCK"
 BASE COUNT 343 a 184 c 205 g 282 t
 ORIGIN
 Query Match 62.2%; Score 974.8; DB 5; Length 1014;
 Best Local Similarity 99.8%; Pred. No. 4.8e-201;
 Matches 976; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 589 ATTAGAGCTTACGACCAAAACCAACATTTATGAGGATCTAGAAAAGTTAGGCTG 648
 |||||||
 Db 13 ATTGAGCTTACGACCAAAACCAACATTTATGAGGATCTAGAAAAGTTAGGCTG 72
 QY 649 GAACAACTTACTGGTCAATGTTCTTCAAGTTTGAAGATTTGTTAGTTAGTCAAGAT 708
 |||||||
 Db 73 GAACAACTTACTGGTCAATGTTCTTCAAGTTTGAAGATTTGTTAGTTAGTCAAGAT 132
 QY 709 AAAAATTAATATTTCCATTAATATGATCCAGAAATTCGGAGTTTGGCAAT 768
 |||||||
 Db 133 AAAAATTAATATTTCCATTAATATGATCCAGAAATTCGGAGTTTGGCAAT 192
 QY 769 AGATATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828
 |||||||
 Db 193 AGATATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252
 QY 829 GATACATTTCTTAAAGCGGCTCATTTATGACAACTGCAAAATGGCATCAAGCAGTA 888
 |||||||
 Db 253 GATACATTTCTTAAAGCGGCTCATTTATGACAACTGCAAAATGGCATCAAGCAGTA 312
 QY 889 AAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 |||||||
 Db 313 AAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
 QY 949 ATGATTTCTTTAACCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
 |||||||
 Db 373 ATGATTTCTTTAACCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
 QY 1009 TCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
 |||||||
 Db 433 TCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
 QY 1069 GCCGAATTAAGATTTATTCAGTTTCAAGCGGAATTAATAGCATCTGCTAGTAGT 1128
 |||||||
 Db 493 GCCGAATTAAGATTTATTCAGTTTCAAGCGGAATTAATAGCATCTGCTAGTAGT 552
 QY 1129 GGCACATTAATATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
 |||||||
 Db 553 GGCACATTAATATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 612
 QY 1189 ACAGATGAAGATTTTAAAGCGGAGAGTACAAAATTTCTCGAGAAAATGGCTCAA 1248
 |||||||
 Db 613 ACAGATGAAGATTTTAAAGCGGAGAGTACAAAATTTCTCGAGAAAATGGCTCAA 672
 QY 1249 ACACCAATTCAGTGTGATGAGGACGAGAAAATAGTCTGATTAAGAGCTTCTTGGA 1308
 |||||||

Db 673 ACCACATTCAGTGGATGGAGCGAGAAAAATAGTCTCGATAAAGCACTTCTTGA 732
Oy 1309 AGTGAATTAAGAACCGGGCGTTGGTAACTGAAAACTCACTCTATATAA 1368
Db 733 AGTGAGATTAAGAACCGGGCGTTGGTAACTGAAAACTCACTCTATATAA 792
Oy 1369 GATAAATAATGATATATCTACCTTTCACACACCTGCTGGATTAAGTCCAGCGCTCAAC 1428
Db 793 GATAAATAATGATATATCTACCTTTCACACACCTGCTGGATTAAGTCCAGCGCTCAAC 852
Oy 1429 GACTTGGTTAGCAAAAAACAACACTGCTGTGATATTAATCACTGTTTAACTCAGCT 1488
Db 853 GACTTGGTTAGCAAAAAACAACACTGCTGTGATATTAATCACTGTTTAACTCAGCT 912
Oy 1489 ATTGAAGCACTGAACCGTTTCAATGAAATATGATCAATGATCAACGCTGCTAGT 1548
Db 913 ATTGAAGCACTGAACCGTTTCAATGAAATATGATCAATGATCAACGCTGCTAGT 972
Oy 1549 GACACGCTGTGTAATGA 1566
Db 973 GACACGCTGTGTAATGA 990

RESULT 10
LOCUS A56795 1014 bp DNA
DEFINITION Sequence 3 from Patent WO9628551.
ACCESSION A56795
VERSION A56795.1 GI:3712810
KEYWORDS
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Tlball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,
Bennett and Alice,M.
TITLE VACCINES FOR PLAGUE
JOURNAL Patent: WO 9628551-A 3 19-SEP-1996;
COMMENT SECR DEFENCE (GB) 2A 9602036 960716
Other publication AU 4951196 961002.
FEATURES
source location/Qualifiers
1..1014
/organism="Yersinia pestis"
/db_xref="taxon:532"
CDS
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAA03420.1"
/db_xref="GI:3712811"
/translation="GIPGIRAYEONPOHFIEDLEKRVLEQJLTHGSSVLELYLVKQ
KNIDISIKYPRKDSFVAFNVRVTDIELEKLAIFYLEPAALIKGHHYDQNLGKIK
RVKELSSPTOMELAFMAVMEFSLADIDDDILKVIYDSNHHGDARSKEEL
AELTAELKIVSVIOAEIKHLSSTGCTINIHKSINMDKNKXTGDEIDFASAEYKI
LEKRPQTTIOVDSSEKIVSIKDFLGSNKRITGALGNKNSIYNKDNNEISHATTTC
SDKSRPNDLVSTQSLDTSIFSNFSIEALNRFIOKYDSVMORLDDISGK"

BASE COUNT 343 a 184 c 205 g 282 t
ORIGIN

Query Match 62.2%; Score 974.8; DB 5; Length 1014;
Best Local Similarity 99.8%; Pred. No. 4.8e-201;
Matches 976; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 589 ATTAGAGCTACGACAAACCAACACATTATTAGATCTAGAAATAAGTAGGGG 648
Db 13 ATTGAGCTACGACAAACCAACACATTATTAGATCTAGAAATAAGTAGGGG 72
Oy 649 GACACGCTGTGTAATGA 1566
Db 73 GACACGCTGTGTAATGA 990

Oy 709 AAAATATAGATATTTCCATTAAATATGATCCGAAAAAGATGGAGTTTGGCAAT 768
Db 133 AAAATATAGATATTTCCATTAAATATGATCCGAAAAAGATGGAGTTTGGCAAT 192
Oy 769 AGAGTAATTAAGTATGATATGAAATGCTCAAGAAATCTAGCTTATTTCTACCCGAG 828
Db 193 AGAGTAATTAAGTATGATATGAAATGCTCAAGAAATCTAGCTTATTTCTACCCGAG 252
Oy 829 GATACATTTCTTAAAGCGGCTATATATGACAAACACTGCAAAATGGCATCAAGGAGTA 888
Db 253 GATGCAATTTCTTAAAGCGGCTATATATGACAAACACTGCAAAATGGCATCAAGGAGTA 312
Oy 889 AAAGATCTCTGATATATCCCGGAATACAAAGGGAATGGCGGCTTATGCGACATA 948
Db 313 AAAGATCTCTGATATATCCCGGAATACAAAGGGAATGGCGGCTTATGCGACATA 372
Oy 949 ATGATTTCTCTTTTAAACCGCGATCGTATCGATGATGATTTTGAAGTATTTGAT 1008
Db 373 ATGATTTCTCTTTTAAACCGCGATCGTATCGATGATGATTTTGAAGTATTTGAT 432
Oy 1009 TCAATGAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
Db 433 TCAATGAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
Oy 1069 GCCGAATTAAGATTTATTTCAAGTTTCAAGCCGAAATTAATAGCATCTGCTAGT 1128
Db 493 GCCGAATTAAGATTTATTTCAAGTTTCAAGCCGAAATTAATAGCATCTGCTAGT 552
Oy 1129 GGCACATTAATATTCATGATTAATATTCATGATGATTAATTAATTAATTAATTAAT 1188
Db 553 GGCACATTAATATTCATGATTAATATTCATGATGATTAATTAATTAATTAATTAAT 612
Oy 1189 ACAGATGAAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTGAGAAAAATGCTCAA 1248
Db 613 ACAGATGAAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTGAGAAAAATGCTCAA 672
Oy 1249 ACCACATTCAGTGGATGGAGCGAGAAAAATAGTCTGATTAAGACTTTCTTGA 1308
Db 673 ACCACATTCAGTGGATGGAGCGAGAAAAATAGTCTGATTAAGACTTTCTTGA 732
Oy 1309 AGTGAATTAAGAACCGGGCGTTGGTAACTGAAAACTCACTCTATATAA 1368
Db 733 AGTGAGATTAAGAACCGGGCGTTGGTAACTGAAAACTCACTCTATATAA 792
Oy 1369 GATAAATAATGATATATCTACCTTTCACACACCTGCTGGATTAAGTCCAGCGCTCAAC 1428
Db 793 GATAAATAATGATATATCTACCTTTCACACACCTGCTGGATTAAGTCCAGCGCTCAAC 852
Oy 1429 GACTTGGTTAGCAAAAAACAACACTGCTGTGATATTAATCACTGTTTAACTCAGCT 1488
Db 853 GACTTGGTTAGCAAAAAACAACACTGCTGTGATATTAATCACTGTTTAACTCAGCT 912
Oy 1489 ATTGAAGCACTGAACCGTTTCAATGAAATATGATCAATGATCAACGCTGCTAGT 1548
Db 913 ATTGAAGCACTGAACCGTTTCAATGAAATATGATCAATGATCAACGCTGCTAGT 972
Oy 1549 GACACGCTGTGTAATGA 1566
Db 973 GACACGCTGTGTAATGA 990

RESULT 11
AF167310 981 bp DNA BCT 17-APR-2000
LOCUS Yersinia pestis strain Angola V antigen (lcrV) gene, complete cds.
ACCESSION AF167310
VERSION AF167310.1 GI:7578514
KEYWORDS
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.

FEATURES	Location/Qualifiers
1	001

1. .981

1. .981

1. .981

Db	481	ACCGCGAATTAAAGATTATTATTCAGTATTATCAAGCGCAAAATTAAATAGACATCTGCTACT	540
Qy	1126	AGTGCACCATTAATATCATCATATTAATTCATTAATCTCTCATGGATTAATAATTTATATGCT	1185
Db	541	AATGGCACCATTAATATCATCATATTAATTCATTAATCTCTATGGATTAATAATTTATATGCT	600
Qy	1186	TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTCGAGAAATAGCT	1245
Db	601	TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTCGAGAAATAGCT	660
Qy	1246	CAAACCCACATTCAGGTGGATGGAGCGAGAGAAAAAATATATCTCGATTAAGACATTTCTT	1305
Db	661	CAAACCCACATTCAGGTGGATGGAGCGAGAGAAAAAATATATCTCGATTAAGACATTTCTT	720
Qy	1306	GGAAGTGAATAAAAAAGAACCGGGGCGTTGGTAAATCTGAAAACTCATCTATATAT	1365
Db	721	GGAAGTGAATAAAAAAGAACCGGGGCGTTGGTAAATCTGAAAACTCATCTATATAT	780
Qy	1366	AAAGATAATTAATGAATTAATCTCTCATTTGGCACACCTGCTCGATTAAGCCAGGCGCTC	1425
Db	781	AAAGATAATTAATGAATTAATCTCTCATTTGGCACACACAGCTGGGATTAAGTCCAGGCGCTC	840
Qy	1426	AACGACTGGTTAGCCAAAAAAACAATCAAGCTGCTGATTAATTAATCAATCACGTTTAATTCA	1485
Db	841	AACGACTGGTTAGCCAAAAAAACAATCAAGCTGCTGATTAATTAATCAATCACGTTTAATTCA	900
Qy	1486	GCTATTGAAGCACTACACCGTTTCATTCAGAAATATGATTCAGTGATGCAACGCTCGCTA	1545
Db	901	GCTATTGAAGCACTACACCGTTTCATTCAGAAATATGATTCAGTGATGCAACGCTCGCTA	960
Qy	1546	GATGACACGCTGTGTAATGA	1566
Db	961	GATGACACGCTGTGTAATGA	981

Matches 977; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 1 ATGATTAGAGCCTACGAACAAACCACACATTTTATTGAGGATCTAGAAATGTTAGG 60

DD 01 GGGAAACACCTACCTGGTCAATGGTCTTCAGTTTATGAGAGAAATGGTTCAGTTAGTCAAA 120

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
[
\
]
^
_
`
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x
y
z
{
|
}
~
_

CERTIFICATE OF ADOPTION

046 **ՀԱՅԿԱՍՏԱՆԻ ՀԱՆՐԱՊԵՏՈՒԹՅԱՆ ՎՃԱՌԱԿԱՆ ԱՄՈՑՈՂՈՒԹՅԱՆ ԳԼԽԱՎՈՐ ԴԻՐԱՋԻ ՎԵՐԱԶՈՐՈՒՄԸ:**

ov 1006 GATTCAATGATCATCATGTGATGCCCGTAGCAAGTTGCCGTGAAGAAATTACCTGAAGCTT 1066

QY 1066 ACCGCCGATTAAGATTTATTCAGTTATTCAGCCGAATTATAAGCATCTGCTAGT 112

.....

LOCUS	AF167309	975 bp	DNA	BCT	17-APR-2000
-------	----------	--------	-----	-----	-------------

VERSION AF167309.1 GI:7578512

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia

TITLE	FIKELANDER, A.M. and KELLIN, P.
Diversity in a variable-number tandem repeat from <i>Yersinia pestis</i>	

AUTHORS H111, K.K. and Jackson, P.J.

FEATURES

Location/Qualifiers

```

/db_xref="taxon:632"

```

```
/gene="IcrV"
```

```
/protein_id=AA64076.1"
/db_xref="GI:7578513"
```

/TRANSLATION- MIKAYEQNPQHFJEDLEKVRVEQLIGHSSVLEELVQLVKDKN


```

Db 181 AATAGATATTACTGATATATCGATTGCTCAGAAAAATCTACTATTATTTCACCC 240
Oy 826 GAGATACCATCTTAAAGCGGCTCATTTATGACAACTGCAAAATGGCATCAAGCA 885
Db 241 GAGATGCCATCTTAAAGCGGCTCATTTATGACAACTGCAAAATGGCATCAAGCA 300
Oy 886 GTAAAGAGCTCTTGAATCATCGCCGAATACAAATGGGAATGGCGGCTCATGCA 945
Db 301 GTAAAGAGCTCTTGAATCATCGCCGAATACAAATGGGAATGGCGGCTCATGCA 360
Oy 946 GTATGATCTCTCTTAAAGCGGCTCATTTATGACAACTGCAAAATGGCATCAAG 1005
Db 361 GTATGATCTCTCTTAAAGCGGCTCATTTATGACAACTGCAAAATGGCATCAAG 420
Oy 1006 GATTCATGATCATCATGATGATGCGCCGATAGCAAGTTGCGTGAAGAAATTAAGT 1065
Db 421 GATTCATGATCATCATGATGATGCGCCGATAGCAAGTTGCGTGAAGAAATTAAGT 480
Oy 1066 ACCGCCGAATTAAGATTTATTCAGTTATTCAGCCGAATTAATAGCATCTGCTAGT 1125
Db 481 ACCGCCGAATTAAGATTTATTCAGTTATTCAGCCGAATTAATAGCATCTGCTAGT 540
Oy 1126 AGTGGCAACATTAATATTCATGATTAATATTCATGATTAATATTAATGAT 1185
Db 541 GGTGGCAACATTAATATTCATGATTAATATTCATGATTAATATTAATGAT 600
Oy 1186 TATACAGATGAAGATTTTAAAGCGGCGAGATCAAAATTTCTCGAGAAATGCT 1245
Db 601 TATACAGATGAAGATTTTAAAGCGGCGAGATCAAAATTTCTCGAGAAATGCT 660
Oy 1246 CAACCCACCATTCAGTGGATGGAGCGAGAAAAAATAGTCTCGATTAAGACCTTTCT 1305
Db 661 CAACCCACCATTCAGTGGATGGAGCGAGAAAAAATAGTCTCGATTAAGACCTTTCT 720
Oy 1306 GGAATGAGAAATTAAGAAACCGGCGGCTGGTAAATCTAAATACATCACTTATAT 1365
Db 721 GGAATGAGAAATTAAGAAACCGGCGGCTGGTAAATCTAAATACATCACTTATAT 780
Oy 1366 AAGATTAATGAATTAATCTACATTTGCCACACCTGCTCGATTAAGTCAAGCGGCTC 1425
Db 781 AAGATTAATGAATTAATCTACATTTGCCACACCTGCTCGATTAAGTCAAGCGGCTC 840
Oy 1426 AAGCACTTGTTAGCCAAAAAACAACACTGCTGCTGATTAATACATCACTTTAATCA 1485
Db 841 AAGCACTTGTTAGCCAAAAAACAACACTGCTGCTGATTAATACATCACTTTAATCA 900
Oy 1486 GCTATTGAAGCACTGAAACGCTTCAATGAAATATGATTCAGTGAAGCTGCTCA 1545
Db 901 GCTATTGAAGCACTGAAACGCTTCAATGAAATATGATTCAGTGAAGCTGCTCA 960
Oy 1546 GATGACACCTCTGTAAATGA 1566
Db 961 GATGACACCTCTGTAAATGA 981

RESULT 14
YEPICRGVHP 2201 bp DNA BCT 26-APR-1993
LOCUS Yersinia pseudotuberculosis V-antigen (lcrG, lcrV, lcrH) genes,
DEFINITION complete cds.
ACCESSION M57893
VERSION M57893.1 GI:155456
KEYWORDS V-antigen.
SOURCE Yersinia pseudotuberculosis (strain YPIII (PIBI) DNA.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 2201)
AUTHORS Bergman,T., Hakansson,S., Forsberg,A., Norlander,L., Macellaro,A.,
Baechman,A., Boelijn,I. and Wolf-Watz,H.
TITLE Analysis of the V antigen lcrvH-yopD operon of Yersinia
pseudotuberculosis: Evidence for a regulatory role of lcrH and lcrV

```

```

JOURNAL J. Bacteriol. 173, 1607-1616 (1991)
MEDLINE 91154114
FEATURES
SOURCE
Location/Qualifiers
1..2201
/organism="Yersinia pseudotuberculosis"
/strain="YPIII (PIBI)"
/db_xref="taxon:633"
219..224
/gene="lcr operon"
219..258
/gene="lcr operon"
-35..signal
-10..signal
RBS
gene
264..551
/gene="lcrG"
264..551
CDS
/gene="lcrG"
/codon_start=1
/transl_table=1
/evidence=experimental
/product="V-antigen"
/protein_id="AA127644.1"
/db_xref="GI:153457"
/transl_table="MKSHFDEYDKTLQAEALVDSHRAKLIQENCADIGLTPPAV
MKIPAGSAEIKPAERELDEIKRQERQPOHPHYDGKRRPKPTMGQII"
553..1533
/gene="lcrV"
553..1533
CDS
/gene="lcrV"
/codon_start=1
/transl_table=1
/evidence=experimental
/product="V-antigen"
/protein_id="AA127645.1"
/db_xref="GI:153458"
/transl_table="MIRAYBONPOHFIEDLEKRVEDQTLGHSSVLELYVOLYKKNI
DISIKYPRDSEVFANRYVITDDIELKILAYLPDAILKGHYDNOXONGIKRYK
EFLSSNTOWELAEFMAVYHFSLTADRIDDLKIVDSMNHHGDRSLRELAEL
TAEIKYISVIOAEIKNLSSGCTINIDKSLNMDKLYGTDEITKAAEYKILEK
MPQTTIDBGETEKIYVSIKPNLEBKRTGALGKDSYINKNLNSHFAITCSOK
SRPNDLVSOCTTQSDITSFNSAIEALNRFQKXDSVMORLLDTSGR"
1546..2052
/gene="lcrH"
1546..2052
CDS
/gene="lcrH"
/codon_start=1
/transl_table=1
/evidence=experimental
/product="V-antigen"
/protein_id="AA127646.1"
/db_xref="GI:155459"
/transl_table="MOEETDIOEQIAMSFLKGGTITAMLEISDTLEIOLYSIAF
NOYSGKEDAHKVFQALCVLDHYDSFFGLACQKQAMQYDLAHSYSYQIMIK
EPREPFAACELQKGLAESEGLFIAOELIADKPEKELSTRVSSMLEPARKREM
EHECVDPNP"
BASE COUNT 714 a 450 c 485 g 552 t
ORIGIN

Query Match 60.9%; Score 953.8; DB 1; Length 2201;
Best Local Similarity 98.3%; Pred. No. 1.6e-196;
Matches 964; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Oy 586 ATGATTAGGCTTACGACAAACCCACACATTTATGAGATCTAGAAAAAGTAGG 645
Db 553 ATGATTAGGCTTACGACAAACCCACACATTTATGAGATCTAGAAAAAGTAGG 612
Oy 646 GTGGAACACTTACGCGCATGCTTCTCACTTTAGAGAAATGGTCACTAGTCAAA 705
Db 613 GTGGAACACTTACGCGCATGCTTCTCACTTTAGAGAAATGGTCACTAGTCAAA 672

```

OY	706	GATAAATAATGTAATTTCCTTCAATTAATATGATCCAGAAAAGATTCGAGGTTTTGGC	765
Db	673	GATAAAAATATATGATTTTCCATTAAATATATGATCCAGAAAAGATTCGAGGTTTTGGC	732
OY	766	AATAGAGTAATTACTGATGATATCGAATTGCTCAAGAAAATCTAGCTTATTTTACC	825
Db	733	AATAGAGTAATTACTGATGATATCGAATTGCTCAAGAAAATCTAGCTTATTTTACC	792
OY	826	GAGGATACCATTTCTTAAAGCGGGTGATTTATGACAACCAACTGCATAATGGCATCAAGGA	885
Db	793	GAGGATGCCATTCTTAAAGCGGGTGATTTATGACAACCAACTGCATAATGGCATCAAGGA	852
OY	886	GTPAAAAGTTCCCTGGAATCATGGCCGGAATACACAATGGGAATTCGGGGCGTCAATGGCA	945
Db	853	GTPAAAAGTTCCCTGGAATCATGGCCGGAATACACAATGGGAATTCGGGGCGTCAATGGCA	912
OY	946	GTAATGCATTTCTCTTTAAACCGCCGATGTATGATGATATTTGAAAATGATTTGT	1005
Db	913	GTAATGCATTTCTCTTTAAACCGCCGATGTATGATGATATTTGAAAATGATTTGT	972
OY	1006	GATTCAATGATCATCATGATGATGATGCCCCGTGACCAAGTTGCGTGAAGAATTTAGCTGAGCTT	1065
Db	973	GATTCAATGATCATCATGATGATGATGCCCCGTGACCAAGTTGCGTGAAGAATTTAGCTGAGCTT	1032
OY	1086	ACCGCGGAATTTAAATATTTATTCAGTTTTCAGACCGCAAGTAATTAATAGCATGTGCTAGT	1125
Db	1033	ACCGCGGAATTTAAATATTTATTCAGTTTTCAGACCGCAAGTAATTAATAGCATGTGCTAGT	1092
OY	1126	AGTGCACCATTAATATTCATGATTAATATTCATATTCATGATGATTAATTAATATGAT	1185
Db	1093	AGTGCACCATTAATATTCATGATTAATATTCATATTCATGATGATTAATTAATATGAT	1152
OY	1186	TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACCAAATTTCTGAGAAAATGCT	1245
Db	1153	TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACCAAATTTCTGAGAAAATGCT	1212
OY	1246	CAAACCCACATTCAGAGTGGAGGAGCGAGAAAATAATATGATGCTCATTAAGACCTTTCTT	1305
Db	1213	CAAACCCACATTCAGAGTGGAGGAGCGAGAAAATAATATGATGCTCATTAAGACCTTTCTT	1272
OY	1306	GGAATGTGAATTAANAAGAACCGGGCGTTGGGTATGTCGAAAACCTCATCTATATAT	1365
Db	1273	GGAATGTGAATTAANAAGAACCGGGCGTTGGGTATGTCGAAAACCTCATCTATATAT	1332
OY	1366	AAAGATTAATTAATTAATTCCTCACCTTTGGCACACACCTGCTCGATAAGTCCAGGCCGCTC	1425
Db	1333	AAAGATTAATTAATTAATTCCTCACCTTTGGCACACACCTGCTCGATAAGTCCAGGCCGCTC	1392
OY	1426	AACGACTGGTTAGCCAAAACCAACTCAGCTGTGTGATATTACATCACGTTTAAATCA	1485
Db	1393	AACGACTGGTTAGCCAAAACCAACTCAGCTGTGTGATATTACATCACGTTTAAATCA	1452
OY	1486	GCTATTGAAGCCTGACCGTTGATTCGAAATATGATTCGATATGCAACGCTGGCTA	1545
Db	1453	GCTATTGAAGCCTGACCGTTGATTCGAAATATGATTCGATATGCAACGCTGGCTA	1512
OY	1546	GATGACAGCTCTGTGTAATGA	1566
Db	1513	GATGACAGCTCTGTGTAATGA	1533
RESULT	15		
AF102990/c			
LOCUS	AF102990	69673 bp	DNA circular BCT 18-MAY-1999
DEFINITION	Yersinia enterocolitica plasmid pYve227, complete sequence.		
ACCESSION	AF102990	AF054978 AF054979 AF054980 AF054981 AF080156 Z69926 L06216	
	AF080155 AF020264 AF050104 AF054977 AF033863 U02499 U08019 U21297		
VERSION	M22781 U08222 Z18539 U94827		
KEYWORDS	AF102990.1 GI:4324323		
SOURCE	Yersinia enterocolitica.		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		

REFERENCE
AUTHORS
TITLE

Yersinia.
1 (bases 32056 to 32871)
Cornelis,G., Sulters,C., de Rouvrol,C.L. and Michiels,T.
Homology between virF, the transcriptional activator of the
Yersinia virulence regulon, and ArcA, the Escherichia coli
arabinose operon regulator
J. Bacteriol. 171 (1), 254-262 (1989)

JOURNAL
MEDLINE
REFERENCE
AUTHORS

2 (bases 15074 to 15994)
Hakansson,S., Bergman,T., Vanooteghem,J.C., Cornelis,G. and
Wolf-Watz,H.
YopB and YopD constitute a novel class of Yersinia Yop proteins
Infect. Immun. 61 (1), 71-80 (1993)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE

3 (bases 45325 to 45717)
Wattiau,P. and Cornelis,G.R.
SYE, a chaperone-like protein of Yersinia enterocolitica involved
in the secretion of YopE
Mol. Microbiol. 8 (1), 123-131 (1993)

JOURNAL
MEDLINE
REFERENCE
AUTHORS

4 (bases 23950 to 25269)
Woestyn,S., Allaoui,A., Wattiau,P. and Cornelis,G.R.
YscN, the putative energizer of the Yersinia Yop secretion
machinery
J. Bacteriol. 176 (6), 1561-1569 (1994)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE

5 (bases 29897 to 30961)
Allaoui,A., Woestyn,S., Sulters,C. and Cornelis,G.R.
YscU, a Yersinia enterocolitica inner membrane protein involved in
Yop secretion
J. Bacteriol. 176 (15), 4534-4542 (1994)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE

6 (bases 43292 to 43717)
Wattiau,P., Bernier,B., Deslee,P., Michiels,T. and Cornelis,G.R.
Individual chaperones required for Yop secretion by Yersinia
Proc. Natl. Acad. Sci. U.S.A. 91 (22), 10493-10497 (1994)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE

7 (bases 31537 to 31932)
Allaoui,A., Scheen,R., Lambert de Rouvrol,C. and Cornelis,G.R.
VirG, a Yersinia enterocolitica lipoprotein involved in Ca²⁺
dependency, is related to exsB of Pseudomonas aeruginosa
J. Bacteriol. 177 (15), 4230-4237 (1995)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE

8 (bases 53135 to 53497)
Stainier,I., Iriarte,M. and Cornelis,G.R.
YscM and YscM2, two Yersinia enterocolitica proteins causing
downregulation of yop transcription
Mol. Microbiol. 26 (4), 833-843 (1997)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE

9 (bases 18681 to 18968)
Boyd,A.P., Sory,M.P., Iriarte,M. and Cornelis,G.R.
Heparin interferes with translocation of Yop proteins into HeLa
cells and binds to LcrE, a regulatory component of the Yersinia yop
apparatus
Mol. Microbiol. 27 (2), 425-436 (1998)

JOURNAL
MEDLINE
REFERENCE
AUTHORS

10 (bases 22612 to 22890)
Iriarte,M., Sory,M.P., Boland,A., Boyd,A.P., Mills,S.D.,
Lambertmont,I. and Cornelis,G.R.
YyeA, a protein involved in control of Yop release and in
translocation of Yersinia Yop effectors
EMBO J. 17 (7), 1907-1918 (1998)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE

11 (bases 1 to 69673)
Iriarte,M., Lambertmont,I., Kerbourch,C. and Cornelis,G.R.
Detailed genetic map of the pYve27 plasmid of Yersinia
enterocolitica serotype O:9
Unpublished

JOURNAL
REFERENCE
AUTHORS
TITLE

12 (bases 1 to 69673)
Iriarte,M., Lambertmont,I., Kerbourch,C. and Cornelis,G.R.
Direct Submission
Submitted (30-Oct-1998) Microbial Pathogenesis Unit, Christian de
Duke Institute of Cellular Pathology and Faculté de Médecine.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 14:59:32 ; Search time 66.74 Seconds

(Without alignments)
5870.554 Million cell updates/sec

Title: US-08-699-716a-1
Perfect score: 1566
Sequence: 1 ATGGCCATCATCATCATCA.....ATGACACGTCTGTGAATCA 1566

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapept 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1472	94.0	1530	1	T38249 Y. pestis Fl/V ant
2	1409	90.0	1462	1	T38256 Y. pestis Fl/V ant
3	976.8	62.4	1014	1	T04222 Partial lcrv (V an
4	975.2	62.3	1014	1	T38242 Y. pestis lcrv (V
5	974.8	62.2	1014	1	T38243 Y. pestis lcrv (V
6	973.2	62.1	1014	1	T04223 Partial lcrv (V an
7	517.2	33.0	544	1	V41594 Nucleotide sequenc
8	516.8	33.0	544	1	O92819 Yersinia pestis ca
9	516.8	33.0	544	1	V41596 Nucleotide sequenc
10	510	32.6	510	1	T38248 Y. pestis Fl antig
11	452.8	28.9	474	1	V41595 Nucleotide sequenc
12	452.8	28.9	474	1	V41600 Nucleotide sequenc
13	452.2	28.9	541	1	O92817 Yersinia pestis ca
14	452.2	28.9	541	1	T38244 Y. pestis cafl (Fl
15	451	28.8	542	1	O92818 Yersinia pestis ca
16	448	28.6	450	1	V41601 Nucleotide sequenc
17	447	28.5	447	1	V41609 Nucleotide sequenc
18	402	25.7	576	1	V41598 Nucleotide sequenc
19	339	21.6	513	1	V41599 Nucleotide sequenc
20	68.8	4.4	861	1	O81500 SV anti-rev sequenc
21	68.8	4.4	861	1	T45347 Single chain sry a
22	68.4	4.4	375	1	V54056 DNA fragment 2 of
23	68.4	4.4	375	1	V54057 DNA fragment 3 of
24	68.4	4.4	375	1	V54058 DNA fragment 4 of
25	54	3.4	1546	1	V30575 Clostridium botuli
26	53.4	3.4	1278	1	V41263 Chlamydomonas rein
27	53.4	3.4	1402	1	T29246 Type A neurotoxin
28	53.4	3.4	1402	1	T30572 Clostridium botuli
29	53.4	3.4	1460	1	V30583 Clostridium botuli
30	53.4	3.4	1463	1	V30584 Clostridium botuli
31	53.4	3.4	1469	1	V30591 Clostridium botuli
32	53.4	3.4	1472	1	V30585 Clostridium botuli
33	53.4	3.4	1502	1	V30588 Clostridium botuli

34	53.4	3.4	1526	1	V30580 Clostridium botuli
35	53.4	3.4	1535	1	V30596 Clostridium botuli
36	53.4	3.4	1547	1	V30581 Clostridium botuli
37	46.8	3.0	104	1	T33359 Ich-2 gene 5' prim
38	45	2.9	21591	1	X13047 Enterococcus faeca
39	42	2.7	110000	1	X20248_01 Continuation (2 of
40	42	2.7	110000	1	X20248_02 Continuation (3 of
41	40	2.6	110000	1	V21309_15 Continuation (16 of
42	39	2.5	4435	1	O29686 CCV-c54 spike gene
43	38.8	2.5	1116	1	O52273 Soybean seed stora
44	38.6	2.5	1124	1	T29508 CONSTANS gene. Iso
45	38	2.4	7953	1	V74463 Staphylococcus aur

ALIGNMENTS

RESULT 1
ID T38249
AC T38249: standard; DNA; 1530 BP.
DT 28-DEC-1996 (first entry)
DE Y. pestis Fl/V antigen gene fusion.
KW Plaque; vaccine; genetic immunisation; V antigen; lcrv;
KW Fl antigen; cafl; ds.
OS Chimeric Yersinia pestis strain GB;
FH Key
FT cds
FT 13.1518
FT Location/Qualifiers
FT /*tag- a
FT /product- Fl/V fusion protein
FT signal_peptide 13.75
FT /*tag- b
FT mat_peptide 76.1515
FT /*tag- c
FT 1522
FT mutation
FT /*tag- d
FT /*note- "base T at position 1522 was modified
FT from G to create a second in-frame stop
FT codon"

FT WO9628551-A1.
FT 19-SEP-1996.
FT 13-MAR-1996: G00571.
FT 13-MAR-1995: GB-005059.
FT 15-SEP-1995: GB-018946.
FT 05-DEC-1995: GB-024825.
FT (MINA) UK SEC FOR DEFENCE.
FT Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
FT WPI: 96-433824/43.
FT P-PSDB: W01044.
FT Yersinia pestis V antigen and Fl antigen or their protective
FT PT Claim 41; Page 65-69; 98pp; English.
FT A gene fusion (T38249) comprises coding sequences for the Yersinia
CC pestis Fl antigen (see also T38248) including the signal sequence
CC and for the Y. pestis V antigen (see also T38243), joined by a
CC sequence encoding a 6-amino acid peptide linker. It was obt'd. by
CC PCR amplification of the individual genes using primers (see also
CC T38245) based on the genes and including linker sequences. The
CC gene fusion (see also T38256) can be used to produce Fl/V fusion
CC protein (W01044) in transformed cells, esp. gut-colonising
CC organisms, to induce an immune response against Y. pestis, the
CC causative organism of plague.
SQ Sequence 1530 BP; 499 A; 296 C; 312 G; 423 T;

Query Match 94.0%; Score 1472; DB 1; Length 1530;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 149; Conservative 0; Mismatches 5; Indels 9; Gaps 1;
OY 68 ATATGAAAAAATCAGTTCCTATTCGCATTCGATATTTGGAACATTCGACATCTGTA 127
DB 11 ATATGAAAAAATCAGTTCCTATTCGCATTCGATATTTGGAACATTCGACATCTGTA 70

```

OY 128 ATGGGAGATTAACTGCAAGCAGCAGTCAAGCACTCTTGTGAACGACCCGCA 187
    |||
DB 71 ATGGGAGATTAACTGCAAGCAGCAGTCAAGCACTCTTGTGAACGACCCGCA 130
OY 188 TCACCTTTCAATTAAGGAAGGCGCTCAATTGCAATTTGCAATGGAACATCGATA 247
    |||
DB 131 TCACCTTTCAATTAAGGAAGGCGCTCAATTGCAATTTGCAATGGAACATCGATA 190
OY 248 CAGAATTACTTGTGTAGCTACTCTTGGGCGCTATAAAACAGGACCACTAGACAT 307
    |||
DB 191 CAGAATTACTTGTGTAGCTACTCTTGGGCGCTATAAAACAGGACCACTAGACAT 250
OY 308 CTGTAACTTACAGATGCGGGGATGCCATGTAATTAACATTTACTTTCAGAGATG 367
    |||
DB 251 CTGTAACTTACAGATGCGGGGATGCCATGTAATTAACATTTACTTTCAGAGATG 310
OY 368 GAAATTAACCACTACTACTCAAAAAGTATGCAAGAGATTCAGAGATTTGATATCT 427
    |||
DB 311 GAAATTAACCACTACTACTCAAAAAGTATGCAAGAGATTCAGAGATTTGATATCT 370
OY 428 CTCTAAGGTAAGCGGTGAGAACCTTGTGGGGATGACGCTGTTGGGTACGGGAGCC 487
    |||
DB 371 CTCTAAGGTAAGCGGTGAGAACCTTGTGGGGATGACGCTGTTGGGTACGGGAGCC 430
OY 488 AGGATTTCTTGTGCTCAATTTGTTCCAAAAGCGGTAACTTGCAAGAGTAATACA 547
    |||
DB 431 AGGATTTCTTGTGCTCAATTTGTTCCAAAAGCGGTAACTTGCAAGAGTAATACA 490
OY 548 CTGATGCTTAACCGTAACCGTATCTTAACCAAGATTCAT-----GATTAGGCT 598
    |||
DB 491 CTGATGCTTAACCGTAACCGTATCTTAACCAAGATTCAT-----GATTAGGCT 550
OY 599 ACGAACAACCAACCAACATTTTATGAGATCTAGAAAAGTTAGGTGGAACAATTA 658
    |||
DB 551 ACGAACAACCAACCAACATTTTATGAGATCTAGAAAAGTTAGGTGGAACAATTA 610
OY 659 CTGTGATGATTTCTGATTTAGAAAAGTTGTTGATGTAAGTCAAAAGATTAATAG 718
    |||
DB 611 CTGTGATGATTTCTGATTTAGAAAAGTTGTTGATGTAAGTCAAAAGATTAATAG 670
OY 719 AATTTTCCATTAATTAATGATCCCAAGAAAAGTTGCGAGTTTGGCAATAGATTAAT 778
    |||
DB 671 AATTTTCCATTAATTAATGATCCCAAGAAAAGTTGCGAGTTTGGCAATAGATTAAT 730
OY 779 CTGATGATTAATGATGATCAAGAAAAGTTGCTAATTTTACCCGAGGATACCATTC 838
    |||
DB 731 CTGATGATTAATGATGATCAAGAAAAGTTGCTAATTTTACCCGAGGATACCATTC 790
OY 839 TTAAGGCGGTCATTAATGACCAACCACTGCAAAATGCGATCAAGCGATTAAGAGTTCC 898
    |||
DB 791 TTAAGGCGGTCATTAATGACCAACCACTGCAAAATGCGATCAAGCGATTAAGAGTTCC 850
OY 899 TTGAATCATGCGGATTAACCAAGTGGGATGCGGCGTTCAAGGATGCAATTTCT 958
    |||
DB 851 TTGAATCATGCGGATTAACCAAGTGGGATGCGGCGTTCAAGGATGCAATTTCT 910
OY 959 CTTTAAACCCGATGATGATGATGATTTGAAAAGTATGTTGATTAAGTAATG 1018
    |||
DB 911 CTTTAAACCCGATGATGATGATGATTTGAAAAGTATGTTGATTAAGTAATG 970
OY 1019 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078
    |||
DB 971 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1030
OY 1079 AATTTTATCAATTAATTAATGAGCGGAAATTAATTAAGATCTGCTAGATGCGACATTA 1138
    |||
DB 1031 AATTTTATCAATTAATTAATGAGCGGAAATTAATTAAGATCTGCTAGATGCGACATTA 1090
OY 1139 AATTCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1198
    |||
DB 1091 AATTCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1150

```

```

OY 1199 AGATTTTAAAGCCAGCGAGATACAAATTTCTGAGAAAATGCTCAAAACCATTC 1258
    |||
DB 1151 AGATTTTAAAGCCAGCGAGATACAAATTTCTGAGAAAATGCTCAAAACCATTC 1210
OY 1259 AGGTGATGCGGAGCGAAGAAAATATGCTGATTAAGACATTTTGTGAAGTGAATA 1318
    |||
DB 1211 AGGTGATGCGGAGCGAAGAAAATATGCTGATTAAGACATTTTGTGAAGTGAATA 1270
OY 1319 AAGAAGCGGGCGTGTGGTATCTGAAAACCTCTACTCTATTAATTAAGATTAATG 1378
    |||
DB 1271 AAGAAGCGGGCGTGTGGTATCTGAAAACCTCTACTCTATTAATTAAGATTAATG 1330
OY 1379 AATTAATCTCACTTTGCAACCACTGCTGGAATTAATCCAGCGGCTCAACGATTTGTTA 1438
    |||
DB 1331 AATTAATCTCACTTTGCAACCACTGCTGGAATTAATCCAGCGGCTCAACGATTTGTTA 1390
OY 1439 GCCAAAACCACTGATGCTGATTAATTAATCACTGTTTAATTCAGCTTAATGAAGC 1498
    |||
DB 1391 GCCAAAACCACTGATGCTGATTAATTAATCACTGTTTAATTCAGCTTAATGAAGC 1450
OY 1499 TGAACCGTTTCAATGAAATTAATGATGATGATGATGATGATGATGATGATGATG 1558
    |||
DB 1451 TGAACCGTTTCAATGAAATTAATGATGATGATGATGATGATGATGATGATGATG 1510
OY 1559 GTAATGA 1566
    |||
DB 1511 GTAATGA 1518

RESULT 2
T38256
ID T38256 standard; DNA; 1462 BP.
AC T38256;
DT 28-DEC-1996 (first entry)
DE Y. pestis F1/V antigen gene fusion.
KW Plaque; vaccine; genetic immunisation; V antigen; lcrV;
KW F1 antigen; catl; ds.
OS Chimeric Yersinia pestis strain GB;
OS Chimeric synthetic.
FH key location/qualifiers
FT cds 8..1450
FT /*tag- a
FT /*product- F1/V fusion protein
FT 452..472
FT misc_feature
FT /*tag- b
FT /*note- "bases 452-472 is a sequence derived
FT from PCR primers"
FN W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MIRA ) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tilball RM, Williamson ED;
DR WPI: 96-433824/43.
DR P-PedsB; W01045.
PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts- useful in vaccine for protection against plague
PS Disclosure: Page 51-55; 98pp; English.
CC A gene fusion (T38256) comprises coding sequences for the Yersinia
CC pestis F1 antigen (see also T38244) (without the signal sequence)
CC and for the Y. pestis V antigen (see also T38243), joined by a
CC sequence encoding a 6-amino acid peptide linker. It was obtd. by
CC PCR amplification of the individual genes using primers (see also
CC T38245) based on the genes and including linker sequences. The
CC gene fusion (see also T38249) can be used to produce F1/V fusion
CC protein (W01045) in transformed cells, esp. gut-colonising
CC organisms, to induce an immune response against Y. pestis, the
CC causative organism of plague.
SQ Sequence 1462 BP; 476 A; 285 C; 301 G; 400 T;

```


Query Match 90.0%; Score 1409; DB 1; Length 1462;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1431; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

131 CGGCATTTAACTGCAAGCACACGCAAGGCACTCTGTTGAACCGCCCGCATCA 190
 |||||
 6 CGGCATTTAACTGCAAGCACACGCAAGGCACTCTGTTGAACCGCCCGCATCA 65

191 CTCTTACATATAAGAGAGCGCTCATATTAATATGACATGGAACATCGATACAG 250
 |||||
 66 CTCTTACATATAAGAGAGCGCTCATATTAATATGACATGGAACATCGATACAG 125

251 AATTACTGTTGGTACGCTTACTCTTGGCGGCTATAAAGAGAGACCATGACATCTG 310
 |||||
 126 AATTACTGTTGGTACGCTTACTCTTGGCGGCTATAAAGAGAGACCATGACATCTG 185

311 TTAACCTTACAGATGCGCGGGGCGGATCCCATGACTTAACTTACTTCTCAGAGTGA 370
 |||||
 186 TTAACCTTACAGATGCGCGGGGCGGATCCCATGACTTAACTTACTTCTCAGAGTGA 245

371 ATACACCAATCTACTACAAAAGATGAGCAAGGATTTAGAGATTTGATATCTCTC 430
 |||||
 246 ATACACCAATCTACTACAAAAGATGAGCAAGGATTTGATATCTCTC 305

431 CTAGAGTAAACGGGTGAACCTTGTGGGGATGACGTCGCTTGGCTACGCGCAGCAG 490
 |||||
 306 CTAGAGTAAACGGGTGAACCTTGTGGGGATGACGTCGCTTGGCTACGCGCAGCAG 365

491 ATTCTCTTGTGCTCAATTTGGTCCAAAGCGGTAACTTGCAGAGGTAAATACACTG 550
 |||||
 366 ATTCTCTTGTGCTCAATTTGGTCCAAAGCGGTAACTTGCAGAGGTAAATACACTG 425

551 ATGCTGTAAACCGGTACCGTATCTAACCAAGATTCAT-----GATTAGAGCCTACG 601
 |||||
 426 ATGCTGTAAACCGGTACCGTATCTAACCAAGATTCATCGAGGTCGTATTAGCCTACG 485

602 AACCAAAACCCACAACTTTTATGAGATCTAGAAAAAGTTAGGTTGGAACCACTTACTG 661
 |||||
 486 AACCAAAACCCACAACTTTTATGAGATCTAGAAAAAGTTAGGTTGGAACCACTTACTG 545

662 GTCATGCTCTTCAGATTGTAAGAGATTTGGTTCAGTTAGCAAGATTAATATAGATA 721
 |||||
 546 GTCATGCTCTTCAGATTGTAAGAGATTTGGTTCAGTTAGCAAGATTAATATAGATA 605

722 TTTCATTAATATGATCCAGAAAAGATTCGAGGTTTTCGATAGAGTAAATACAG 781
 |||||
 606 TTTCATTAATATGATCCAGAAAAGATTCGAGGTTTTCGATAGAGTAAATACAG 665

782 ATGATATGCAATGCTCAAGAAAATCTAGCTTATTTTCTACCGAGGATACCATTTCTTA 841
 |||||
 666 ATGATATGCAATGCTCAAGAAAATCTAGCTTATTTTCTACCGAGGATACCATTTCTTA 725

842 AAGCGCGCTATTATGCAACCAACTGCCAAATGCAATCAAGCAGATTAAGAGTCTCTTG 901
 |||||
 726 AAGCGCGCTATTATGCAACCAACTGCCAAATGCAATCAAGCAGATTAAGAGTCTCTTG 785

902 AATCATCGCGAATACATGGAATGCGGGGCTTCAGAGGATTAAGATTCATTTCTTCT 961
 |||||
 786 AATCATCGCGAATACATGGAATGCGGGGCTTCAGAGGATTAAGATTCATTTCTTCT 845

962 TAACCGCCGATCGATATGATGATATTTTGAAGTATGTTGATTCATGATATCATC 1021
 |||||
 846 TAACCGCCGATCGATATGATGATATTTTGAAGTATGTTGATTCATGATATCATC 905

1022 ATGATATGCAATGCTCAAGAAAATCTAGCTTATTTTCTACCGAGGATACCATTTCTTA 1081
 |||||
 906 ATGATATGCAATGCTCAAGAAAATCTAGCTTATTTTCTACCGAGGATACCATTTCTTA 965

1082 TTATTCAGTATTAAGCGGAATTAATTAAGATCTGCTAGTAGGACATTAATAA 1141
 |||||
 966 TTATTCAGTATTAAGCGGAATTAATTAAGATCTGCTAGTAGGACATTAATAA 1025

1142 TCATGATTAATTCATTAATTCATGATTAATAATTAATTAATGTTATACAGATGAAGAGA 1201

Query Match 62.4%; Score 976.8; DB 1; Length 1014;
 Best Local Similarity 99.8%; Pred. No. 7.2e-251;
 Matches 978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1026 TCCATGATTAATTCATTAATTCATGATTAATAATTTATATGTTATACAGATGAAGAGA 1085
 |||||

1202 TTTTAAAGCCGCGAGAGTACAAAATTTCTGAGAAAAATGCTTCAACACCATTCAG 1261
 |||||

1086 TTTTAAAGCCGCGAGAGTACAAAATTTCTGAGAAAAATGCTTCAACACCATTCAG 1145
 |||||

1262 TGGATGGAGCAGAAAAAATAGTCTGATTAAGAGATTTCTTGAAGTGAATAA 1321
 |||||

1146 TGGATGGAGCAGAAAAAATAGTCTGATTAAGAGATTTCTTGAAGTGAATAA 1205
 |||||

1322 GAACGGGCGTGGGTAACTGAAAAACTACTCTTATATAAAGATTAATGAAT 1381
 |||||

1206 GAACGGGCGTGGGTAACTGAAAAACTACTCTTATATAAAGATTAATGAAT 1265
 |||||

1382 TATCTACCTTGGCAGACCTGCTGGATTAAGTCCAGGCGCTCAACGACTTGTTAGCC 1441
 |||||

1266 TATCTACCTTGGCAGACCTGCTGGATTAAGTCCAGGCGCTCAACGACTTGTTAGCC 1325
 |||||

1442 AAAAAACAAGTCTGCTGATTAATCAATCAAGTTTAAATCAGCTATGAAGCCTGA 1501
 |||||

1326 AAAAAACAAGTCTGCTGATTAATCAATCAAGTTTAAATCAGCTATGAAGCCTGA 1385
 |||||

1502 ACCGTTTCAATCAGAAATATGATTCAGTATGATCAACGCTCTAGATGACAGCTTGATA 1561
 |||||

1386 ACCGTTTCAATCAGAAATATGATTCAGTATGATCAACGCTCTAGATGACAGCTTGATA 1445
 |||||

1562 AATGA 1566
 |||||

1446 AATGA 1450

RESULT 3
 T04222
 ID T04222 standard; DNA; 1014 BP.
 AC T04222;
 DT 18-APR-1996 (first entry).
 DE Partial IcrV (V antigen) gene of Y. pestis.
 KW IcrV; V antigen; virulence; plague; vaccine; epitope; ss.
 OS Yersinia pestis.
 FH key
 FT cds
 FT 1..990
 FT /*tag= a
 FT /note= "V antigen"

W09524475-21.
 PD 14-SEP-1995.
 PF 06-MAR-1995; G00481.
 PR 08-MAR-1994; GB-004577.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Leary SEC, Titball RW, Williamson ED, Leary SE;
 DR WPI: 95-328268/42.
 DR P-PsDB: R/9961.
 PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
 oral or parenteral vaccines for protection against plague
 PS Claim 6; Page 11-13; 25pp; English.
 CC T04222-23 are DNA sequences (IcrV) encoding all or a protective epitopic
 part of the mature V protein of Yersinia pestis. The protein was
 expressed as a fusion protein with maltose binding protein or
 glutathione-S-transferase in 3 different plasmid vectors. Y. pestis is
 the highly virulent causative organism of plague in a wide range of
 CC animals, including man. The V antigen (IcrV) is an unstable 37.3 kDa
 monomeric peptide encoded on the ca. 70 kb Icr plasmid. The V antigen
 is postulated to act as a virulence antigen, and transformed
 CC microorganisms contg. recombinant DNA encoding a V antigen protein/
 CC peptide are useful in vaccines to protect against plague.
 SQ Sequence 1014 BP; 346 A; 181 C; 201 G; 286 T;

Query Match 62.4%; Score 976.8; DB 1; Length 1014;
 Best Local Similarity 99.8%; Pred. No. 7.2e-251;
 Matches 978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

587 TGATGACCTCAGCAACAAACCAACATTTATATGAGATCTAGAAAAAGTTAGCG 646

```

Db 11 TCATTAGAGCCTACGACAAACCAACCAACCTTTATTAGAGATCTAGAAAAAATTGAGG 70
QY 647 TGAACAACCTACTGTCATGCTGTTCTTCAGTTTGAAGAATGGTTCAGTTAGTCAAG 706
Db 71 TGAACAACCTACTGTCATGCTGTTCTTCAGTTTGAAGAATGGTTCAGTTAGTCAAG 130
QY 707 ATAAAAATATAGATATTTCCATTAAATATGATCCAGAAAAAGATTCGAGGTTTTGGCA 766
Db 131 ATAAAAATATAGATATTTCCATTAAATATGATCCAGAAAAAGATTCGAGGTTTTGGCA 190
QY 767 ATAGAGATATCTGATGATATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 826
Db 191 ATAGAGATATCTGATGATATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 250
QY 827 AGATACCATCTTAAAGGCGGTATATGATGACCAACCACTGCAAAATGGCATCAAGCGAG 886
Db 251 AGATACCATCTTAAAGGCGGTATATGATGACCAACCACTGCAAAATGGCATCAAGCGAG 310
QY 887 TAAAGAGTTCCTGATTCATCGCCGGAATACACAATGGGAATTCGCGGCGTTTCATGGCAG 946
Db 311 TAAAGAGTTCCTGATTCATCGCCGGAATACACAATGGGAATTCGCGGCGTTTCATGGCAG 370
QY 947 TAATGATTTCTCTTAAACCCCATGCTATGATGATGATGATGATGATGATGATGATG 1006
Db 371 TAATGATTTCTCTTAAACCCCATGCTATGATGATGATGATGATGATGATGATGATGATG 430
QY 1007 ATTAAGATATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
Db 431 ATTAAGATATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490
QY 1067 CCGCGCAATTAAGATTTATTCAGTTATTCAGCCGGAATTAATAGCATCTGCTAGTA 1126
Db 491 CCGCGCAATTAAGATTTATTCAGTTATTCAGCCGGAATTAATAGCATCTGCTAGTA 550
QY 1127 GTGCGACCATTAATATTCATATTAATCCATTATTCATGATGATGATGATGATGATGATG 1186
Db 551 GTGCGACCATTAATATTCATATTAATCCATTATTCATGATGATGATGATGATGATGATGATG 610
QY 1187 ATACAGATGAAGATTTTAAAGCCAGCGAGATGACCAAAATTCCTGAGAAATGCTC 1246
Db 611 ATACAGATGAAGATTTTAAAGCCAGCGAGATGACCAAAATTCCTGAGAAATGCTC 670
QY 1247 AAACCCACCATTCAGGTGATGAGGAGCGAGAAAAAATAGTCTCGATTAAGGACTTCTG 1306
Db 671 AAACCCACCATTCAGGTGATGAGGAGCGAGAAAAAATAGTCTCGATTAAGGACTTCTG 730
QY 1307 GAATGAGATTAAGAAAGAACCGGGGCGTGGTATCTGAAAAAATCTCATCTCTATATA 1386
Db 731 GAATGAGATTAAGAAAGAACCGGGGCGTGGTATCTGAAAAAATCTCATCTCTATATA 790
QY 1367 AAGATTAATTAAGATTTATTCATCTTGGCACACCATCTGCTGGATTAAGCCAGCGGCTCA 1426
Db 791 AAGATTAATTAAGATTTATTCATCTTGGCACACCATCTGCTGGATTAAGCCAGCGGCTCA 850
QY 1427 ACAGCTGGTTAGCCAAAAAACAACCTAGCTGTCTGATTAATACACAGTTTAAATTCAG 1486
Db 851 ACAGCTGGTTAGCCAAAAAACAACCTAGCTGTCTGATTAATACACAGTTTAAATTCAG 910
QY 1487 CTATTGAAGCCTGAACCGTTTCATTCAGAAATATGATTCGATGATGATGATGATGATG 1546
Db 911 CTATTGAAGCCTGAACCGTTTCATTCAGAAATATGATTCGATGATGATGATGATGATG 970
QY 1547 ATGACAGCTCTGTAATGA 1566
Db 971 ATGACAGCTCTGTAATGA 990

```

RESULT 4
T38242
ID T38242 standard: DNA; 1014 BP.
AC T38242;
DT 28-DEC-1996 (first entry)

```

DE Y. pestis lcrV (V antigen) gene.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW Fl antigen; ds.
OS Yersinia pestis
FH Key location/Qualifiers
FT cds 1..990
FT misc_feature 1..6
FT tag- a
FT tag- b
FT note- "vector pMAL-p2 or pMAL-c2-derived bases"
FT mutation
FT /tag- c
FT /note- "base 1006 is altered to a T to create a
FT second in-frame stop codon"
MO9628551-A1.
PN 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCR, Tibball RM, Williamson ED;
DR WPI; 96-433824/43.
DR P-PDB; W01040.
PT Yersinia pestis V antigen and Fl antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Claim 5; Page 25-28; 98pp; English.
CC A lcrV gene sequence (T38242) codes for the Yersinia pestis V
CC antigen (W01040), which is capable of evoking protective immune
CC responses in animals. The gene was amplified from Y. pestis
CC DNA by PCR using primers (T38250-51) homologous to the 5' and 3'
CC ends of the gene. The gene was inserted into vector pMAL-p2,
CC pMAL-c2 or pGEX-5x-2 (see also T38243) to allow prodn. of
CC recombinant V antigen for use in vaccines against plague.
CC Expression in gut-colonising organisms and attenuated Salmonella
CC typhl allows prodn. of live vaccines. Fl/V antigen fusions were
CC also created (see also T38249 and T38256). The gene can itself be
CC used in genetic vaccines.
SQ Sequence 1014 BP; 347 A; 180 C; 201 G; 286 T;

Query Match 62.3%; Score 975.2; DB 1; Length 1014;
Best Local Similarity 99.7%; Pred. No. 1.9e-250;
Matches 977; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 587 TGAATTAGAGCTTCGAAACCAACCAACCACTTTTATGAGATCTAGAAAAAGTTAGG 646
Db 11 TCATTAGAGCCTACGACAAACCAACCAACCACTTTTATGAGATCTAGAAAAAGTTAGG 70
QY 647 TGAACAACCTACTGTCATGCTGTTCTTCAGTTTGAAGAATGGTTCAGTTAGTCAAG 706
Db 71 TGAACAACCTACTGTCATGCTGTTCTTCAGTTTGAAGAATGGTTCAGTTAGTCAAG 130
QY 707 ATAAAAATATAGATTTTCCATTAAATATGATCCAGAAAAAGATTCGAGGTTTTGGCA 766
Db 131 ATAAAAATATAGATTTTCCATTAAATATGATCCAGAAAAAGATTCGAGGTTTTGGCA 190
QY 767 ATAGAGATATCTGATGATATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 826
Db 191 ATAGAGATATCTGATGATATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 250
QY 827 AGATACCATCTTAAAGGCGGTATATGATGACCAACCACTGCAAAATGGCATCAAGCGAG 886
Db 251 AGATACCATCTTAAAGGCGGTATATGATGACCAACCACTGCAAAATGGCATCAAGCGAG 310
QY 887 TAAAGAGTTCCTGATTCATCGCCGGAATACACAATGGGAATTCGCGGCGTTTCATGGCAG 946
Db 311 TAAAGAGTTCCTGATTCATCGCCGGAATACACAATGGGAATTCGCGGCGTTTCATGGCAG 370
QY 947 TAATGATTTCTCTTAAACCCCATGCTATGATGATGATGATGATGATGATGATGATGATG 1006
Db 371 TAATGATTTCTCTTAAACCCCATGCTATGATGATGATGATGATGATGATGATGATGATG 430

```

QY	1007	ATTCAATGAATCAATCAGGATGCGCCGATGCAAGTGGGTGAAGATTAAGCTGAGCTTA	1066
QY	1007	ATTCAATGAATCAATCAGGATGCGCCGATGCAAGTGGGTGAAGATTAAGCTGAGCTTA	1066
Db	431	ATTCAATGAATCAATCAGGATGCGCCGATGCAAGTGGGTGAAGATTAAGCTGAGCTTA	490
QY	1067	CCGCGGATTAAGATTTATTCAGTATTCACGCCCAATTAATAAGCATCTGCTAGTA	1126
Db	491	CCGCGGATTAAGATTTATTCAGTATTCACGCCCAATTAATAAGCATCTGCTAGTA	550
QY	1127	GTGGCACCATTAAATATTCATGATTAATTCATTATATCTCATGGATATAAAATTAATGTT	1186
QY	1127	GTGGCACCATTAAATATTCATGATTAATTCATTATATCTCATGGATATAAAATTAATGTT	1186
Db	551	GTGGCACCATTAAATATTCATGATTAATTCATTATATCTCATGGATATAAAATTAATGTT	610
QY	1187	ATACAGATGAAGATTTTAAAGCCAGCCGACAGTACAAATTCCTCGAGAAATGCCTC	1246
Db	611	ATACAGATGAAGATTTTAAAGCCAGCCGACAGTACAAATTCCTCGAGAAATGCCTC	670
QY	1247	AAACACCATTCAGGTGATGAGGAGGAGAAAAAATAGTCTCGATTAAGGACTTCTTG	1306
Db	671	AAACACCATTCAGGTGATGAGGAGGAGAAAAAATAGTCTCGATTAAGGACTTCTTG	730
QY	1307	GAACTGAGAAATAAAAGACCGGGGCGTTGGTATCTGAAAAAAGCTCATCTTTAATA	1366
Db	731	GAACTGAGAAATAAAAGACCGGGGCGTTGGTATCTGAAAAAAGCTCATCTTTAATA	790
QY	1367	AAGTAATTAATGAATTTCTCATCTTTGCCACCACTGCTCGGATTAATGACGCGCTCA	1426
Db	791	AAGTAATTAATGAATTTCTCATCTTTGCCACCACTGCTCGGATTAATGACGCGCTCA	850
QY	1427	ACGACTTGGTTAGCCAAAAAAGACAGTCAAGTCTATATTTATCATACAGTTTATATAG	1486
Db	851	ACGACTTGGTTAGCCAAAAAAGACAGTCAAGTCTATATTTATCATACAGTTTATATAG	910
QY	1487	CTATTGAGACACTGAACCGCTTCATTCAGAAATATGATTCAGTGAATGCAACGCTGCTAG	1546
Db	911	CTATTGAGACACTGAACCGCTTCATTCAGAAATATGATTCAGTGAATGCAACGCTGCTAG	970
QY	1547	ATGACACGCTCTGTTAATGA 1566	
Db	971	ATGACACGCTCTGTTAATGA 990	
RESULT	5		
ID	T38243		
ID	T38243	standard; DNA; 1014 BP.	
AC	T38243		
DT	28-DEC-1996	(first entry)	
DE	X. pestis lcrv (V antigen) gene.		
KW	Plague; vaccine; genetic immunisation; V antigen; lcrv;		
KW	F1 antigen; ds.		
OS	Yersinia pestis strain GB.		
FH	Key	location/Qualifiers	
FT	cds	1..990	
FT		/*tag- a	
FT	misc-feature	1..10	
FT		/*tag- b	
FT	mutation	/note- "vector pGEX-5x-2-derived bases"	
FT		16	
FT		/*tag- c	
FT	mutation	/note- "base 16 is altered to a C from an A to createan EcoRI site"	
FT		1006	
FT		/*tag- d	
FT		/note- "base 1006 is altered to a T to create a second in-frame stop codon"	
PN	W09628551-A1.		
PD	19-SEP-1996.		
PF	13-MAR-1996; G00571.		
PR	13-MAR-1995; GB-005059.		
PR	15-SEP-1995; GB-018946.		
PR	05-DEC-1995; GB-024825.		
PA	(MINA) UK SEC FOR DEFENCE.		
PI	Bennett AM, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;		
PR	WPI: 96-433824/43.		

Query Match	Best Local Similarity	Score	DB 1	Length	1014	DB 2	Length	1014	DB 3	Length	1014	DB 4	Length	1014	DB 5	Length	1014	DB 6	Length	1014	DB 7	Length	1014	DB 8	Length	1014	DB 9	Length	1014	DB 10	Length	1014	DB 11	Length	1014	DB 12	Length	1014	DB 13	Length	1014	DB 14	Length	1014	DB 15	Length	1014	DB 16	Length	1014	DB 17	Length	1014	DB 18	Length	1014	DB 19	Length	1014	DB 20	Length	1014	DB 21	Length	1014	DB 22	Length	1014	DB 23	Length	1014	DB 24	Length	1014	DB 25	Length	1014	DB 26	Length	1014	DB 27	Length	1014	DB 28	Length	1014	DB 29	Length	1014	DB 30	Length	1014	DB 31	Length	1014	DB 32	Length	1014	DB 33	Length	1014	DB 34	Length	1014	DB 35	Length	1014	DB 36	Length	1014	DB 37	Length	1014	DB 38	Length	1014	DB 39	Length	1014	DB 40	Length	1014	DB 41	Length	1014	DB 42	Length	1014	DB 43	Length	1014	DB 44	Length	1014	DB 45	Length	1014	DB 46	Length	1014	DB 47	Length	1014	DB 48	Length	1014	DB 49	Length	1014	DB 50	Length	1014	DB 51	Length	1014	DB 52	Length	1014	DB 53	Length	1014	DB 54	Length	1014	DB 55	Length	1014	DB 56	Length	1014	DB 57	Length	1014	DB 58	Length	1014	DB 59	Length	1014	DB 60	Length	1014	DB 61	Length	1014	DB 62	Length	1014	DB 63	Length	1014	DB 64	Length	1014	DB 65	Length	1014	DB 66	Length	1014	DB 67	Length	1014	DB 68	Length	1014	DB 69	Length	1014	DB 70	Length	1014	DB 71	Length	1014	DB 72	Length	1014	DB 73	Length	1014	DB 74	Length	1014	DB 75	Length	1014	DB 76	Length	1014	DB 77	Length	1014	DB 78	Length	1014	DB 79	Length	1014	DB 80	Length	1014	DB 81	Length	1014	DB 82	Length	1014	DB 83	Length	1014	DB 84	Length	1014	DB 85	Length	1014	DB 86	Length	1014	DB 87	Length	1014	DB 88	Length	1014	DB 89	Length	1014	DB 90	Length	1014	DB 91	Length	1014	DB 92	Length	1014	DB 93	Length	1014	DB 94	Length	1014	DB 95	Length	1014	DB 96	Length	1014	DB 97	Length	1014	DB 98	Length	1014	DB 99	Length	1014	DB 100	Length	1014	DB 101	Length	1014	DB 102	Length	1014	DB 103	Length	1014	DB 104	Length	1014	DB 105	Length	1014	DB 106	Length	1014	DB 107	Length	1014	DB 108	Length	1014	DB 109	Length	1014	DB 110	Length	1014	DB 111	Length	1014	DB 112	Length	1014	DB 113	Length	1014	DB 114	Length	1014	DB 115	Length	1014	DB 116	Length	1014	DB 117	Length	1014	DB 118	Length	1014	DB 119	Length	1014	DB 120	Length	1014	DB 121	Length	1014	DB 122	Length	1014	DB 123	Length	1014	DB 124	Length	1014	DB 125	Length	1014	DB 126	Length	1014	DB 127	Length	1014	DB 128	Length	1014	DB 129	Length	1014	DB 130	Length	1014	DB 131	Length	1014	DB 132	Length	1014	DB 133	Length	1014	DB 134	Length	1014	DB 135	Length	1014	DB 136	Length	1014	DB 137	Length	1014	DB 138	Length	1014	DB 139	Length	1014	DB 140	Length	1014	DB 141	Length	1014	DB 142	Length	1014	DB 143	Length	1014	DB 144	Length	1014	DB 145	Length	1014	DB 146	Length	1014	DB 147	Length	1014	DB 148	Length	1014	DB 149	Length	101
-------------	-----------------------	-------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	-----

```

QY 1369 GATTAATATGAAATTAATCTCACTTGGCCACCCTGCTCGGATAGTCCAGCCGCTCAAC 1428
DB 793 GATAATATATGAAATTAATCTCACTTGGCCACCCTGCTCGGATAGTCCAGCCGCTCAAC 852
QY 1429 GATTGGTTAGCCAAAAAACAACACTGCTGTGATATTTCATCAGCTTTTAATTCAGCT 1488
DB 853 GACTTGGTTAGCCAAAAAACAACACTGCTGTGATATTTCATCAGCTTTTAATTCAGCT 912
QY 1489 ATGAAGACACTGAACCGTTTCATTGAGAAATATGATCAGTATGCAAGCTGCTAGAT 1548
DB 913 ATGAAGACACTGAACCGTTTCATTGAGAAATATGATCAGTATGCAAGCTGCTAGAT 972
QY 1549 GACACGCTGTGTAATGA 1566
DB 973 GACACGCTGTGTAATGA 990

RESULT 6
T04223
ID T04223 standard; DNA; 1014 BP.
AC T04223;
DT 18-APR-1996 (first entry)
DE Partial LcrV (V antigen) gene of Y. pestis.
KW LcrV; V antigen; virulence; plague; vaccine; epitope; ss.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT cds 1..990
FT FT /*tag= a
PN MO9524475-A1.
PD 14-SEP-1995.
PE 06-MAR-1995; G00481.
PR 08-MAR-1994; GB-004577.
PA (MINA ) UK SEC FOR DEFENCE.
PI Leary SEC, Titchell RW, Williamson ED, Leary SE;
DR MPI, 95-328268/42.
P-PSDB; R/79962.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PT oral or parenteral vaccines for protection against plague
PS Claim 6: Page 15-16: 25pp: English.
CC T04222-73 are DNA sequences (LcrV) encoding all or a protective epitopic
CC part of the mature V protein of Yersinia pestis. The protein was
CC expressed as a fusion protein with maltose binding protein. Y. pestis
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis
CC is the highly virulent causative organism of plague in a wide range of
CC animals, including man. The V antigen (LcrV) is an unstable 37.3 kDa
CC monomeric peptide encoded on the ca. 70 kb Lcr plasmid. The V antigen
CC is postulated to act as a virulence antigen, and transformed
CC microorganisms contg. recombinant DNA encoding a V antigen protein/
CC peptide are useful in vaccines to protect against plague.
SQ Sequence 1014 BP; 343 A; 185 C; 204 G; 282 T;

Query Match 62.1%; Score 973.2; DB 1; Length 1014;
Best Local Similarity 99.7%; Pred. No. 6,6e-250;
Matches 975; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 829 GATACATCTTTAAAGCGGCTATTATGACCAACCACTGCAAAATGGCATCAAGCGAGTA 888
DB 253 GATGCCATCTTTAAAGCGGCTATTATGACCAACCACTGCAAAATGGCATCAAGCGAGTA 312
QY 889 AAGAGTCTCTTGAATCATATCCGCCAATACACAAATGGGAATTTGGGGCTTCATGGCAGTA 948
DB 313 AAGAGTCTCTTGAATCATATCCGCCAATACACAAATGGGAATTTGGGGCTTCATGGCAGTA 372
QY 949 ATGATTTCTCTTAAACCGCGCATGCTATGATGATGATGATGATGATGATGATGATGAT 1008
DB 373 ATGATTTCTCTTAAACCGCGCATGCTATGATGATGATGATGATGATGATGATGATGAT 432
QY 1009 TCAATGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
DB 433 TCAATGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
QY 1069 GCGGATTTAAAGTATTATTCAGTATTCAAGCCGGAATTAATAGCATGTGCTAGTACT 1128
DB 493 GCGGATTTAAAGTATTATTCAGTATTCAAGCCGGAATTAATAGCATGTGCTAGTACT 552
QY 1129 GGCACCATTAATATCCATGATTAATCCATTAATCCATGATTAATTAATGATGATGAT 1188
DB 553 GGCACCATTAATATCCATGATTAATCCATTAATCCATGATTAATTAATGATGATGAT 612
QY 1189 ACAGATGAAGAGATTTTAAAGCCAGCAGAGTACAAATTTCTCGAGAAATGGCTCAA 1248
DB 613 ACAGATGAAGAGATTTTAAAGCCAGCAGAGTACAAATTTCTCGAGAAATGGCTCAA 672
QY 1249 ACCACCATTCAGTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1308
DB 673 ACCACCATTCAGTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 732
QY 1309 AGTGAATTAATGAAGACCGGGGGTGGTGAATCTGAATTAATGAAGACCTTCTATATAA 1368
DB 733 AGTGAATTAATGAAGACCGGGGGTGGTGAATCTGAATTAATGAAGACCTTCTATATAA 792
QY 1369 GATTAATATGAAATTAATCTCACTTGGCCACCCTGCTCGGATAGTCCAGCCGCTCAAC 1428
DB 793 GATAATATATGAAATTAATCTCACTTGGCCACCCTGCTCGGATAGTCCAGCCGCTCAAC 852
QY 1429 GATTGGTTAGCCAAAAAACAACACTGCTGTGATATTTCATCAGCTTTTAATTCAGCT 1488
DB 853 GACTTGGTTAGCCAAAAAACAACACTGCTGTGATATTTCATCAGCTTTTAATTCAGCT 912
QY 1489 ATGAAGACACTGAACCGTTTCATTGAGAAATATGATCAGTATGCAAGCTGCTAGAT 1548
DB 913 ATGAAGACACTGAACCGTTTCATTGAGAAATATGATCAGTATGCAAGCTGCTAGAT 972
QY 1549 GACACGCTGTGTAATGA 1566
DB 973 GACACGCTGTGTAATGA 990

RESULT 7
V41594
ID V41594 standard; DNA; 544 BP.
AC V41594;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of FI antigen nYPI(a)sec544.
KW FI antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT cds 17..529
FT FT /*tag= a
FT FT /product= "FI antigen"
PN WO9824912-A2.
PD 11-JUN-1998.
PE 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haimes EJ, Osorio JE, Thomas RE;
DR MPI, 98-333331/29.
P-PSDB; W59782.

```

PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Page 51; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 544 BP; 164 A; 120 C; 114 G; 146 T;

Query Match 33.0%; Score 517.2; DB 1; Length 544;
 Best Local Similarity 98.5%; Pred. No. 1.4e-128;
 Matches 522; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAGTCCGTTATCGCATTCATGATTTTGGAACTATGCAACTGCTA 127
 DB 15 ATATGAAAAAATCAGTCCGTTATCGCATTCATGATTTTGGAACTATGCAACTGCTA 74
 QY 128 ATGGGAGATTATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 187
 DB 75 ATGGGAGATTATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 134
 QY 188 TCACCTTACATATAGAAAGCGCTCCATTTACATTTATGCAATGGAACATCGATA 247
 DB 135 TCACCTTACATATAGAAAGCGCTCCATTTACATTTATGCAATGGAACATCGATA 194
 QY 248 CAGATTACTTGTGATGACCTTCTTGGCGGCTATTAAGCAAGCAAGCAAGCAAGCA 307
 DB 195 CAGATTACTTGTGATGACCTTCTTGGCGGCTATTAAGCAAGCAAGCAAGCAAGCA 254
 QY 308 CTGTAACTTACAGATCGCGGGGTATCCCATGATTTACTTACTTCTCAGATG 367
 DB 255 CTGTAACTTACAGATCGCGGGGTATCCCATGATTTACTTACTTCTCAGATG 314
 QY 368 GAAATAACCAACCAATTAACAAAGATGATGGCAAGATTTAGAGATTTATATCT 427
 DB 315 GAAATAACCAACCAATTAACAAAGATGATGGCAAGATTTAGAGATTTATATCT 374
 QY 428 CTCTTAAGTAAAGGATGAGACCTTGTGGGATGACGTGCTTGGCTACGGGACGC 487
 DB 375 CTCTTAAGTAAAGGATGAGACCTTGTGGGATGACGTGCTTGGCTACGGGACGC 434
 QY 488 AGGATTTCTTGTGCTCAATGATTCCTCAAGCGGTAACCTTGCAGCAGTAATACA 547
 DB 435 AGGATTTCTTGTGCTCAATGATTCCTCAAGCGGTAACCTTGCAGCAGTAATACA 494
 QY 548 CTGATGCTGTAACCGTATCTAATCAAGATTCATGATTAAGCC 597
 DB 495 CTGATGCTGTAACCGTATCTAATCAAGATTCATGATTAAGCC 544

RESULT 8
 ID 092819
 AC 092819;
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis cafi (F1) antigen in plasmid pFORF1b.
 KW Yersinia pestis; Yersinia pestis; Yersinia pestis typhi;
 KM Yersinia pestis; Yersinia pestis; Yersinia pestis typhi;
 OS Yersinia pestis.
 FH Key
 FT cds
 FT 2-7
 FT /tag- a
 FT /note- "first protein encoded by pFORF1b"
 FT misc-feature
 FT 1.6
 FT /tag- b
 FT misc-feature
 FT 536..541
 FT /tag- c
 FT /note- "cafi open reading frame downstream seq."
 FT 21..530
 FT /tag- d
 FT /note- "cafi fusion protein"

PN WC9518231-A1.
 PD 06-JUL-1995.
 PE 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tildall RW, Williamson ED;
 DR WPI; 95-246396/32.
 DR P-SDS; R76528.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure; Page 19-20; 27pp; English.
 CC The sequence represents the plasmid pFORF1b including the entire
 CC Y. pestis cafi (F1) antigen gene having a 5' tail including a SacI
 CC restriction site, and up to 1400 bp downstream of the cafi ORF.
 CC The DNA construct can be used to transform human or animal gut
 CC colonizing microorganisms, specifically attenuated Salmonella
 CC typhimurium or Yersinia pestis. The transformed microorganisms
 CC can be used as live/attenuated vaccines which induce immune
 CC responses at mucosal surfaces. The vaccines provide protection
 CC against infection with Y. pestis, and are parenterally and orally
 CC active vaccines offering protection against bubonic and pneumonic
 CC plague.
 SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

Query Match 33.0%; Score 516.8; DB 1; Length 542;
 Best Local Similarity 99.0%; Pred. No. 1.8e-128;
 Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAGTCCGTTATCGCATTCATGATTTTGGAACTATGCAACTGCTA 127
 DB 19 ATATGAAAAAATCAGTCCGTTATCGCATTCATGATTTTGGAACTATGCAACTGCTA 78
 QY 128 ATGGGAGATTATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 187
 DB 79 ATGGGAGATTATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 138
 QY 188 TCACCTTACATATAGAAAGCGCTCCATTTACATTTATGCAATGGAACATCGATA 247
 DB 139 TCACCTTACATATAGAAAGCGCTCCATTTACATTTATGCAATGGAACATCGATA 198
 QY 248 CAGATTACTTGTGATGACCTTCTTGGCGGCTATTAAGCAAGCAAGCAAGCAAGCA 307
 DB 199 CAGATTACTTGTGATGACCTTCTTGGCGGCTATTAAGCAAGCAAGCAAGCAAGCA 258
 QY 308 CTGTAACTTACAGATCGCGGGGTATCCCATGATTTACTTACTTCTCAGATG 367
 DB 259 CTGTAACTTACAGATCGCGGGGTATCCCATGATTTACTTACTTCTCAGATG 318
 QY 368 GAAATAACCAACCAATTAACAAAGATGATGGCAAGATTTAGAGATTTATATCT 427
 DB 319 GAAATAACCAACCAATTAACAAAGATGATGGCAAGATTTAGAGATTTATATCT 378
 QY 428 CTCTTAAGTAAAGGATGAGACCTTGTGGGATGACGTGCTTGGCTACGGGACGC 487
 DB 379 CTCTTAAGTAAAGGATGAGACCTTGTGGGATGACGTGCTTGGCTACGGGACGC 438
 QY 488 AGGATTTCTTGTGCTCAATGATTCCTCAAGCGGTAACCTTGCAGCAGTAATACA 547
 DB 439 AGGATTTCTTGTGCTCAATGATTCCTCAAGCGGTAACCTTGCAGCAGTAATACA 498
 QY 548 CTGATGCTGTAACCGTATCTAATCAAGATTCATGATTAAGCC 597
 DB 499 CTGATGCTGTAACCGTATCTAATCAAGATTCATGATTAAGCC 538

RESULT 9
 ID V41596
 AC V41596;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nyf1sec510.

KW Fl antigen; plasmid; vaccine; plague; ds.
 OS Yersinia pestis. Location/Qualifiers
 FH Key 17..532
 FT CDS /*tag= a
 FT /product= "Fl antigen"
 FN WO9824912-A2.
 PD 11-JUN-1998.
 PE 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HEK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI:96-33331/29.
 DR P-PSDB: W59783.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8: Pages 53-54; 75pp: English.
 CC This is the nucleotide sequence of a Yersinia pestis Fl antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 CC Sequence 544 BP; 166 A; 118 C; 112 G; 148 T;
 SQ

Query Match 33.0%; Score 516.8; DB 1; Length 544;
 Best Local Similarity 99.6%; Pred. No. 1.8e-128;
 Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAAGTCCGTTATCGCATTTATTTGGAACTATTCGACCTGCTA 127
 DB 15 ATATGAAAAAATCAAGTCCGTTATCGCATTTATTTGGAACTATTCGACCTGCTA 74

QY 128 ATGGCGAGATTAACTGCAAGCAGCAGTCAAGGCACTCTGTGTAACGAGCCGCA 187
 DB 75 ATGGCGAGATTAACTGCAAGCAGCAGTCAAGGCACTCTGTGTAACGAGCCGCA 134

QY 188 TCACCTTACATATAGAGAGGCGCTCCATTACATTTATGACAAATGGAACATCGATA 247
 DB 135 TCACCTTACATATAGAGAGGCGCTCCATTACATTTATGACAAATGGAACATCGATA 194

QY 248 CAGATTACTTTGTGTAAGCTTACTCTGGGGGCTATTAACAGAACCACTAGACAT 307
 DB 195 CAGATTACTTTGTGTAAGCTTACTCTGGGGGCTATTAACAGAACCACTAGACAT 254

QY 308 CTGTAACTTTACAGATGCGGGGATGCCATGACTTAACTTACTTCTCAGAGATG 367
 DB 255 CTGTAACTTTACAGATGCGGGGATGCCATGACTTAACTTACTTCTCAGAGATG 314

QY 368 GAAATTAACCAACATTAACATAAGATGATGGCAAGGATCTAGAGATTTGATATCT 427
 DB 315 GAAATTAACCAACATTAACATAAGATGATGGCAAGGATCTAGAGATTTGATATCT 374

QY 428 CTCCTAAGTTAAAGGTGAGAACTTGTGGGGATGAGCTCTTGGCTAGGGGAGCC 487
 DB 375 CTCCTAAGTTAAAGGTGAGAACTTGTGGGGATGAGCTCTTGGCTAGGGGAGCC 434

QY 488 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACTTGCAGAGGTAATACA 547
 DB 435 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACTTGCAGAGGTAATACA 494

QY 548 CTGATGCTGTAAACCGTAAACGATATCTAACCAAGATTCAT 587
 DB 495 CTGATGCTGTAAACCGTAAACGATATCTAACCAAGATTCAT 534

RESULT 10
 T38248 standard; DNA; 547 BP.
 ID T38248;
 AC T38248;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis Fl antigen cati gene (including signal sequence).
 KW Plague; vaccine; genetic immunisation; Fl antigen; cati;

KW V antigen; ds. strain GB.
 OS Yersinia pestis. Location/Qualifiers
 FH Key 29..91
 FT signal_peptide /*tag= a
 FT mat_peptide 92..538
 FT /*tag= b
 FN WO9628551-A1.
 PD 19-SEP-1996.
 PE 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RM, Williamson ED;
 DR WPI:96-433824/43.
 DR P-PSDB: W01043.
 PT Yersinia pestis V antigen and Fl antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Claim 41; Page 61-62; 98pp: English.
 CC A DNA sequence (T38248) comprises the cati gene, including the
 CC signal sequence, coding for the Fl antigen (W01043) of Yersinia
 CC pestis. It was obtd. by PCR amplification (see also T38257-58)
 CC of Y. pestis DNA. The PCR product was cloned into plasmid pBKCMV
 CC and the resulting plasmid (pFlAB) was used to transform E. coli Nova
 CC Blue. Purified plasmid, when administered by i.m. injection,
 CC induced an immunoglobulin response to Fl in BALB/c mice. Live
 CC vaccines comprising gut colonising organisms transformed with the
 CC cati gene (see also T38244) can be used to protect a host animal
 CC against plague.
 CC Sequence 547 BP; 165 A; 120 C; 115 G; 147 T;
 SQ

Query Match 33.0%; Score 516.8; DB 1; Length 547;
 Best Local Similarity 99.6%; Pred. No. 1.8e-128;
 Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAAGTCCGTTATCGCATTTATTTGGAACTATTCGACCTGCTA 127
 DB 27 ATATGAAAAAATCAAGTCCGTTATCGCATTTATTTGGAACTATTCGACCTGCTA 86

QY 128 ATGGCGAGATTAACTGCAAGCAGCAGTCAAGGCACTCTGTGTAACGAGCCGCA 187
 DB 87 ATGGCGAGATTAACTGCAAGCAGCAGTCAAGGCACTCTGTGTAACGAGCCGCA 146

QY 188 TCACCTTACATATAGAGAGGCGCTCCATTACATTTATGACAAATGGAACATCGATA 247
 DB 147 TCACCTTACATATAGAGAGGCGCTCCATTACATTTATGACAAATGGAACATCGATA 206

QY 248 CAGATTACTTTGTGTAAGCTTACTCTGGGGGCTATTAACAGAACCACTAGACAT 307
 DB 207 CAGATTACTTTGTGTAAGCTTACTCTGGGGGCTATTAACAGAACCACTAGACAT 266

QY 308 CTGTAACTTTACAGATGCGGGGATGCCATGACTTAACTTACTTCTCAGAGATG 367
 DB 267 CTGTAACTTTACAGATGCGGGGATGCCATGACTTAACTTACTTCTCAGAGATG 326

QY 368 GAAATTAACCAACATTAACATAAGATGATGGCAAGGATCTAGAGATTTGATATCT 427
 DB 327 GAAATTAACCAACATTAACATAAGATGATGGCAAGGATCTAGAGATTTGATATCT 386

QY 428 CTCCTAAGTTAAAGGTGAGAACTTGTGGGGATGAGCTCTTGGCTAGGGGAGCC 487
 DB 387 CTCCTAAGTTAAAGGTGAGAACTTGTGGGGATGAGCTCTTGGCTAGGGGAGCC 446

QY 488 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACTTGCAGAGGTAATACA 547
 DB 447 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACTTGCAGAGGTAATACA 506

QY 548 CTGATGCTGTAAACCGTAAACGATATCTAACCAAGATTCAT 587
 DB 507 CTGATGCTGTAAACCGTAAACGATATCTAACCAAGATTCAT 546

RESULT 11

ID V41595 standard; DNA: 510 BP.
AC V41595;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nYpF1(b)sec544.
KW F1 antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
PN M09824912-A2.
PD 11-JUN-1998;
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HEK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-33331/29.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 53; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SO Sequence 510 BP; 153 A; 114 C; 106 G; 137 T;

Query Match 32.6%; Score 510; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.1e-126;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 ATGAAAAATCAGTCCGTTATGCCATTTGATTTGACATTTGCAATGCACTGCTAT 129
DB 1 ATGAAAAATCAGTCCGTTATGCCATTTGATTTGACATTTGCAATGCACTGCTAT 60
OY 130 GCGGAGATTTACTGACAGCACACTGCAAGGCACTCTTTGAAACGACCGGCATC 189
DB 61 GCGGAGATTTACTGACAGCACACTGCAAGGCACTCTTTGAAACGACCGGCATC 120
OY 190 ACTCTTACATATAGGAAGGCGCTCAATTACATTTATGACAATGGAACATCGATACA 249
DB 121 ACTCTTACATATAGGAAGGCGCTCAATTACATTTATGACAATGGAACATCGATACA 180
OY 250 GAATTAATCTGTTGACGCTTACTCTTGGCGGCTTAAACAGGAAACCACTGCACTCT 309
DB 181 GAATTAATCTGTTGACGCTTACTCTTGGCGGCTTAAACAGGAAACCACTGCACTCT 240
OY 310 GTTAACCTTACAGATGCGGCGGATGCCATGATCTTACATTTACTTCTCAGATGGA 369
DB 241 GTTAACCTTACAGATGCGGCGGATGCCATGATCTTACATTTACTTCTCAGATGGA 300
OY 370 AATAACCAACAATTCATACAAAGTGTGCAAGAGATTTAGAGATTTGATATCTCT 429
DB 301 AATAACCAACAATTCATACAAAGTGTGCAAGAGATTTAGAGATTTGATATCTCT 360
OY 430 CCTAAGGTAAGCGGTGAGAACTTGTGGGGATGACGCTGTTGGCTACGCGCAGCAG 489
DB 361 CCTAAGGTAAGCGGTGAGAACTTGTGGGGATGACGCTGTTGGCTACGCGCAGCAG 420
OY 490 GATTTCTTGTGCTCATTGTTCCAAAGGCGGTAAGTCCAGGAGGTAAATACACT 549
DB 421 GATTTCTTGTGCTCATTGTTCCAAAGGCGGTAAGTCCAGGAGGTAAATACACT 480
OY 550 GATGCTGTACCGTATACCGTATACCAA 579
DB 481 GATGCTGTACCGTATACCGTATACCAA 510

RESULT 12
ID V41600 standard; DNA: 474 BP.
AC V41600;
DT 26-OCT-1998 (first entry)

DE Nucleotide sequence of F1 antigen nYpF1mat474.
KW F1 antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 7..459
FT /tag= a
FT /product= "F1 antigen"

PN M09824912-A2.
PD 11-JUN-1998;
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HEK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-33331/29.
DR P-PSDB: W39787.

PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Pages 59-60; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SO Sequence 474 BP; 143 A; 106 C; 99 G; 126 T;

Query Match 28.9%; Score 452.8; DB 1; Length 474;
Best Local Similarity 99.6%; Pred. No. 1.9e-111;
Matches 454; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 132 GGCAGATTTAATCGAACCACCACTGCAAGGCACTCTTGTGACACGCGCATCAC 191
DB 9 GGCAGATTTAATCGAACCACCACTGCAAGGCACTCTTGTGACACGCGCATCAC 68
OY 192 TCTTACATATAGGAAGGCGCTCCATATACATTTATGACAATGGAACATCGATACAGA 251
DB 69 TCTTACATATAGGAAGGCGCTCCATATACATTTATGACAATGGAACATCGATACAGA 128
OY 252 ATTACTTGTGTGACGCTTACTCTTGGCGGCTATTAACAGGAACCACTAGCACATCTGT 311
DB 129 ATTACTTGTGTGACGCTTACTCTTGGCGGCTATTAACAGGAACCACTAGCACATCTGT 188
OY 312 TAACCTTACAGATGCGGCGGATGCCATGATCTTACATTTACTTCTCAGATGGA 371
DB 189 TAACCTTACAGATGCGGCGGATGCCATGATCTTACATTTACTTCTCAGATGGA 248
OY 372 TAACCAACAATTCATACAAAGTGTGCAAGAGATTTGATATCTCTCC 431
DB 249 TAACCAACAATTCATACAAAGTGTGCAAGAGATTTGATATCTCTCC 308
OY 432 TAAGTAAACGGTGAAGACCTTGTGGGGATGACGCTGTTGGCTACGCGCAGCAGGA 491
DB 309 TAAGTAAACGGTGAAGACCTTGTGGGGATGACGCTGTTGGCTACGCGCAGCAGGA 368
OY 492 TTTCTTGTGCTCATTGTTCCAAAGGCGGTAAGTCCAGGAGGTAAATACACTGA 551
DB 369 TTTCTTGTGCTCATTGTTCCAAAGGCGGTAAGTCCAGGAGGTAAATACACTGA 428
OY 552 TGCTGTACCGTATACCGTATACCAAAGTTCAT 587
DB 429 TGCTGTACCGTATACCGTATACCAAAGTTCAT 464

RESULT 13
ID 092817
AC 092817;
DT 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen in plasmid pFGAL2a.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
OS Yersinia pestis. Location/Qualifiers

FT cds 2..454
 FT /*tag- a
 FT misc-feature 1..6
 FT /*tag- b
 FT /note- "lacZ promoter fusion site"
 FT misc-feature 536..541
 FT /*tag- c
 FT /note- "vector pFAL2a bases"
 PN MO9518231-A1.
 PD 06-JUL-1995.
 PE 23-DEC-1994: G02818.
 PR 24-DEC-1993: GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
 DR WPI: 95-246396/32.
 DR P-PSDB: R76526.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Claim 7: Page 15-16; 27pp; English.
 CC The sequence represents the plasmid pFAL2a construct showing the
 CC fusion of the first few bases of beta-galactosidase in the vector
 CC with the Y. pestis catf (F1) antigen minus its signal sequence and
 CC having a 5' tail including a SacI restriction site, and up to the
 CC catf AAC-3' end with some vector bases. The DNA construct can be
 CC used to transform human or animal gut colonizing microorganisms,
 CC specifically attenuated Salmonella typhimurium or Salmonella typh.
 CC The transformed microorganisms can be used as live/attenuated
 CC vaccines which induce immune responses at mucosal surfaces. The
 CC are parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 541 BP; 163 A; 120 C; 111 G; 147 T;

Query Match 28.9%; Score 452.2; DB 1; Length 541;
 Best Local Similarity 99.3%; Pred. No. 2.8e-111;

Matches 453; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DB 131 CGGAGATTAACTGCAAGCAGCACTGCAAGGCACTCTTGTGAACGACCGCATCA 190
 DB 6 CGGAGATTAACTGCAAGCAGCACTGCAAGGCACTCTTGTGAACGACCGCATCA 65
 QY 191 CTCTACATATAGGAAGCGCTCCCAATTACATTTATGGAATGAAACATGATACAG 250
 DB 66 CTATTACATATAGGAAGCGCTCCCAATTACATTTATGGAATGAAACATGATACAG 125
 QY 251 AATTACTGTTGTAACCTTACTCTTGGCGGCTATTAACAGGACCATGACATCTG 310
 DB 126 AATTACTGTTGTAACCTTACTCTTGGCGGCTATTAACAGGACCATGACATCTG 185
 QY 311 TTAATCTTACAGATGCGCGGGTATCCCATGTAATTACTTACTCTGAGATGAA 370
 DB 186 TTAATCTTACAGATGCGCGGGTATCCCATGTAATTACTTACTCTGAGATGAA 245
 QY 371 ATAAACCAATTCATCAAAAAGTATGCAAGGATTCAGAGATTTGATATCTCTC 430
 DB 246 ATAAACCAATTCATCAAAAAGTATGCAAGGATTCAGAGATTTGATATCTCTC 305
 QY 431 CTAGGTAAGCGTGAGAACCTTGTGGGGATGACGTCGTTGGCTACGCGGACGAG 490
 DB 306 CTAGGTAAGCGTGAGAACCTTGTGGGGATGACGTCGTTGGCTACGCGGACGAG 365
 QY 491 ATTCTTGTGCTCAATGTTCTCAAGGCGGTAACTGACAGAGTAATACATG 550
 DB 366 ATTCTTGTGCTCAATGTTCTCAAGGCGGTAACTGACAGAGTAATACATG 425
 QY 551 ATGCTTAACCGTATCTATCAACAGAAATTCAT 587
 DB 426 ATGCTTAACCGTATCTATCAACAGAAATTCAT 462

RESULT 14

T38244
 ID T38244 standard; DNA; 541 BP.
 AC T38244;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis catf (F1 antigen) gene.
 KW Plague; vaccine; genetic immunisation; F1 antigen; catf;
 KW V antigen; ds.
 OS Yersinia pestis strain GB.
 FH Key location/Qualifiers
 FT misc_recomb 1..6
 FT /*tag- a
 FT /note- "includes the first few bases of the
 FT beta-galactosidase sequence in vector
 FT pFAL2a"
 FT cds 2..457
 FT /*tag- b
 FT /product- mature F1 antigen
 FT misc_recomb 536..541
 FT /*tag- c
 PN MO9628551-A1.
 PD 19-SEP-1996.
 PE 13-MAR-1996: G00571.
 PR 13-MAR-1995: GB-005059.
 PR 15-SEP-1995: GB-018946.
 PR 05-DEC-1995: GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR P-PSDB: W01042.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Claim 7: Page 43-45; 98pp; English.
 CC The catf gene (T38244) codes for the Yersinia pestis F1 antigen
 CC (W01042), which is capable of evoking protective immune responses
 CC in animals. The gene was amplified from Y. pestis DNA by PCR using
 CC primers (T38245-46) homologous to the 5' and 3' ends of the gene.
 CC The catf PCR product was ligated into pUC18 and transformed into E.
 CC coli JM109 to produce vector pFAL2a. The catf gene can be used to
 CC express recombinant F1 antigen for use in vaccines against plague.
 CC Expression in gut-colonising organisms and attenuated Salmonella
 CC typh allows prodn. of live vaccines. The gene can itself be used
 CC in genetic vaccines. F1/V antigen fusions were also created (see
 CC also T38249 and T38256).
 SQ Sequence 541 BP; 163 A; 120 C; 111 G; 147 T;

Query Match 28.9%; Score 452.2; DB 1; Length 541;
 Best Local Similarity 99.3%; Pred. No. 2.8e-111;

Matches 453; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 131 CGGAGATTAACTGCAAGCAGCACTGCAAGGCACTCTTGTGAACGACCGCATCA 190
 DB 6 CGGAGATTAACTGCAAGCAGCACTGCAAGGCACTCTTGTGAACGACCGCATCA 65
 QY 191 CTCTACATATAGGAAGCGCTCCCAATTACATTTATGGAATGAAACATGATACAG 250
 DB 66 CTATTACATATAGGAAGCGCTCCCAATTACATTTATGGAATGAAACATGATACAG 125
 QY 251 AATTACTGTTGTAACCTTACTCTTGGCGGCTATTAACAGGACCATGACATCTG 310
 DB 126 AATTACTGTTGTAACCTTACTCTTGGCGGCTATTAACAGGACCATGACATCTG 185
 QY 311 TTAATCTTACAGATGCGCGGGTATCCCATGTAATTACTTACTCTGAGATGAA 370
 DB 186 TTAATCTTACAGATGCGCGGGTATCCCATGTAATTACTTACTCTGAGATGAA 245
 QY 371 ATAAACCAATTCATCAAAAAGTATGCAAGGATTCAGAGATTTGATATCTCTC 430
 DB 246 ATAAACCAATTCATCAAAAAGTATGCAAGGATTCAGAGATTTGATATCTCTC 305
 QY 431 CTAGGTAAGCGTGAGAACCTTGTGGGGATGACGTCGTTGGCTACGCGGACGAG 490
 DB 306 CTAGGTAAGCGTGAGAACCTTGTGGGGATGACGTCGTTGGCTACGCGGACGAG 365

OY 491 ATTCTTGTTCGCTCAATTGGTTCACAAAGCGGTAACTTCAGCAGTAATACATC 550
 DB 366 ATTTCTTTGTTCGCTCAATTGGTTCACAAAGCGGTAACTTCAGCAGTAATACATC 425
 OY 551 ATGCTGTACCGTAACCGTATCTTAACCAAGAATTCAT 587
 DB 426 ATGCTGTACCGTAACCGTATCTTAACCAAGAATTCAT 462

RESULT 15

O92818 15
 ID Q92818 standard: DNA: 542 BP.

AC Q92818:
 DE 17-DEC-1995 (first entry)
 DT Yersinia pestis cafi (PI) antigen in plasmid pF313a.
 KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 OS Yersinia pestis.
 KW Yersinia pestis.
 OS Yersinia pestis.

Location/Qualifiers
 cds 3..455
 FT /tag- a

FT misc-feature

1..6
 /tag- b
 /note- "E. coli LTB signal peptide bases"

FT misc-feature

536..541
 /tag- c
 /note- "vector pF313a bases"

FT W09518231-A1.

PD 06-JUL-1995.

PE 23-DEC-1994; G02818.

PR 24-DEC-1993; GB-026425.

PA (MINA) UK SEC FOR DEFENCE.

PI Howells A, Leary SPC, Oyston PCF, Tibball RW, Williamson ED;

DR WPI: 95-246396/32.

DR P-PSDB; R/6527.

PT DNA constructs capable of transforming microorganisms - which can be

PT used as live or attenuated vaccines which induce an immune response,

PT against Yersinia pestis, at mucosal surfaces.

PS Disclosure: Page 17-18; 27pp; English.

CC The sequence represents the plasmid pF313a construct showing the

CC fusion of the first few bases of the E. coli LTB signal sequence

CC with the Y. pestis cafi (PI) antigen minus its signal sequence and

CC having a 5' tail including a SacI restriction site, and up to the

CC cafi AAC-3' end with some vector bases. The DNA construct can be

CC used to transform human or animal gut colonizing microorganisms,

CC specifically attenuated Salmonella typhimurium or Salmonella typhi.

CC The transformed microorganisms can be used as live/attenuated

CC vaccines which induce immune responses at mucosal surfaces. The

CC vaccines provide protection against infection with Y. pestis, and

CC are parenterally and orally active vaccines offering protection

CC against bubonic and pneumonic plague.

SO Sequence 342 BP; 163 A; 122 C; 110 G; 147 T;

Query Match 28.8%; Score 451; DB 1; Length 542;

Best Local Similarity 98.9%; Pred. No. 5.9e-11;

Matches 454; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 129 TCGGCGAGATTACTGCAAGCACCAGTGCACAGGCACTCTGTGAACGACCCGCAT 188
 DB 5 TCCCGCAAGATTACTGCAAGCACCAGTGCACAGGCACTCTGTGAACGACCCGCAT 64
 OY 189 CACTTTCATTAAGAGAGGCGCTCCATTAATATATGACAAATGGAACATGCATAC 248
 DB 65 CACTTTCATTAAGAGAGGCGCTCCATTAATATATGACAAATGGAACATGCATAC 124
 OY 249 AGAATTACTTGTGTAGAGCTTACTCTGGGGCTATAAAGAGAACCACTAGACATC 308
 DB 125 AGAATTACTTGTGTAGAGCTTACTCTGGGGCTATAAAGAGAACCACTAGACATC 184
 OY 309 TGTTAACCTTACAGATGCCGCGGTATCCAGTACTTAACATTTACTCTCAGGATG 368

DB 185 TGTTAACCTTACAGATGCCGCGGTATCCAGTACTTAACATTTACTCTCAGGATG 244
 OY 369 AATAACCAACCAATTCACTACAAAGATGAGGAGATCTAGATTTGATATCTC 428
 DB 245 AATAACCAACCAATTCACTACAAAGATGAGGAGATCTAGATTTGATATCTC 304
 OY 429 TCCTAAGGTAAACGAGTGAAGAACCTTGTGGGAGATGACGCTGTGGCTAGCGGACCA 488
 DB 305 TCCTAAGGTAAACGAGTGAAGAACCTTGTGGGAGATGACGCTGTGGCTAGCGGACCA 364
 OY 489 GGATTTCTTTGTGCTGCTCAATTGGTTCACAAAGCGGTAACTTCAGCAGTAATACATC 548
 DB 365 GGATTTCTTTGTGCTGCTCAATTGGTTCACAAAGCGGTAACTTCAGCAGTAATACATC 424
 OY 549 TGATGCTGTAAACCGTAAACCGTATCTTAACCAAGAATTCAT 587
 DB 425 TGATGCTGTAAACCGTAAACCGTATCTTAACCAAGAATTCAT 463

Search completed: August 22, 2000, 15:36:39
 Job time: 2227 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 14:58:41 ; Search time 46.08 Seconds

(without alignments)
4674.749 Million cell updates/sec

Title: US-08-699-716a-1

1566

Sequence: 1 ATGGCCATCATCATCATCA.....ATGACACGCTGTGAATGA 1566

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 243080 segs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0

Maximum DB seq length: 100000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/5C.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/5D.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/PCBUS.COMB.seq:*
7: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1473.6	94.1	1530	4 US-08-913-477-22	Sequence 22, Appl
2	1409	90.0	1462	4 US-08-913-477-16	Sequence 16, Appl
3	976.8	62.4	1014	4 US-08-913-477-1	Sequence 1, Appl
4	974.8	62.2	1014	4 US-08-913-477-3	Sequence 3, Appl
5	513.6	32.8	547	4 US-08-913-477-20	Sequence 20, Appl
6	452.2	28.9	541	4 US-08-913-477-10	Sequence 10, Appl
7	68.4	4.4	375	5 US-08-870-370-7	Sequence 7, Appl
8	68.4	4.4	375	5 US-08-870-370-8	Sequence 8, Appl
9	68.4	4.4	375	5 US-08-870-370-9	Sequence 9, Appl
10	53.4	3.4	1402	2 US-08-480-604A-25	Sequence 25, Appl
11	53.4	3.4	1402	2 US-08-405-966A-25	Sequence 25, Appl
12	48	3.1	7218	1 US-08-232-463-14	Sequence 14, Appl
13	39	2.5	4435	1 US-08-308-872B-5	Sequence 5, Appl
14	38.6	2.5	1124	5 US-08-945-056-1	Sequence 11, Appl
15	37.4	2.4	4365	6 PCT-US91-08525-31	Sequence 31, Appl
16	37.4	2.4	4365	6 PCT-US93-04384-11	Sequence 11, Appl
17	35	2.2	47	4 US-08-600-783-15	Sequence 15, Appl
18	34.8	2.2	1446	4 US-08-560-007B-4	Sequence 4, Appl
19	34.8	2.2	1508	5 US-08-714-918-31	Sequence 31, Appl
20	34.8	2.2	2402	5 US-08-776-265-4	Sequence 4, Appl
21	34.8	2.2	2415	2 US-08-785-428-1	Sequence 1, Appl
22	34.8	2.2	2415	4 US-08-996-797-1	Sequence 1, Appl
23	34.6	2.2	1100	3 US-08-179-537-20	Sequence 20, Appl
24	34	2.2	867	4 US-09-036-582-36	Sequence 36, Appl
25	33.8	2.2	1947	5 US-08-604-991-1	Sequence 1, Appl
26	33.8	2.2	1947	5 US-09-363-639-1	Sequence 1, Appl

27	33.6	2.1	2902	5 US-08-714-918-95	Sequence 95, Appl
28	33.2	2.1	4565	5 US-08-776-265-1	Sequence 1, Appl
29	33.2	2.1	9636	2 US-08-323-170B-1	Sequence 1, Appl
30	33.2	2.1	10564	1 US-08-206-176-5	Sequence 5, Appl
31	33	2.1	1888	2 US-08-229-145-13	Sequence 13, Appl
32	33	2.1	2993	4 US-08-415-593-42	Sequence 42, Appl
33	32.8	2.1	8920	3 US-08-446-855A-1	Sequence 1, Appl
34	32.8	2.1	891	1 US-07-941-523-19	Sequence 19, Appl
35	32.8	2.1	1959	2 US-08-137-175A-1	Sequence 1, Appl
36	32.8	2.1	2369	3 US-08-525-742-1	Sequence 1, Appl
37	32.6	2.1	2653	3 US-08-589-711-1	Sequence 1, Appl
38	32.6	2.1	2794	1 US-07-960-932-1	Sequence 1, Appl
39	32.6	2.1	2794	1 US-07-908-253-1	Sequence 1, Appl
40	32.6	2.1	2794	2 US-08-455-970A-1	Sequence 1, Appl
41	32.6	2.1	2794	2 US-08-187-156-5	Sequence 5, Appl
42	32.6	2.1	2794	3 US-08-694-865-5	Sequence 5, Appl
43	32.6	2.1	2794	4 US-08-878-748-5	Sequence 5, Appl
44	32.6	2.1	2794	4 US-08-535-837-1	Sequence 1, Appl
45	32.6	2.1	2794	5 US-09-124-491-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-913-477-22
Sequence 22, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlingtonton
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 424
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
TELECOMMUNICATION INFORMATION:
REFERENCE/DOCKET NUMBER: 124-599
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4000
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:

LENGTH: 1530 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Yersinia pestis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 13..1515
 US-08-913-477-22

Query Match 94.1% Score 1473.6; DB 4; Length 1530;
 Best Local Similarity 99.1% Pred. No. 0;
 Matches 1495; Conservative 0; Mismatches 4; Indels 9; Gaps 1;

QY 68 AATGTAAGAAATTCAGTCCGTTATGCGCATTTGCAATTTATGGAAGTATGCAATGCTA 127
 DB 11 AATGTAAGAAATTCAGTCCGTTATGCGCATTTGCAATTTATGGAAGTATGCAATGCTA 70
 QY 128 ATGGGCAATTTACTGCAAGCAGCAGTGCAGAGCGCACTCTGTTGAACCGCCGCA 187
 DB 71 ATGGGCAATTTACTGCAAGCAGCAGTGCAGAGCGCACTCTGTTGAACCGCCGCA 130
 QY 188 TCACCTTACATTAAGAAAGCGCTCAATTCATTTATGCAATTTATGGAAGTATGCAATGCTA 247
 DB 131 TCACCTTACATTAAGAAAGCGCTCAATTCATTTATGCAATTTATGGAAGTATGCAATGCTA 190
 QY 248 CAGAAATTTACTGTTGTTAGCTTACTCTTGGCGGTATTAAGAGAACGACATAGACAT 307
 DB 191 CAGAAATTTACTGTTGTTAGCTTACTCTTGGCGGTATTAAGAGAACGACATAGACAT 250
 QY 308 CTGTTAACTTACAGATGCGCGGGGGTATGCCATGTTACTTAACTTACTTCTCAGAGATG 367
 DB 251 CTGTTAACTTACAGATGCGCGGGGGTATGCCATGTTACTTAACTTACTTCTCAGAGATG 310
 QY 368 GAAATTAACCAACATTCATACATAAAGTATGGAAGTATGCAATTTATGGAAGTATGCAATGCTA 427
 DB 311 GAAATTAACCAACATTCATACATAAAGTATGGAAGTATGCAATTTATGGAAGTATGCAATGCTA 370
 QY 428 CTCTTAAGTAAAGCGTGAAGACCTTGGGGGATGACGTCGTTGGCTACGGGAGCC 487
 DB 371 CTCTTAAGTAAAGCGTGAAGACCTTGGGGGATGACGTCGTTGGCTACGGGAGCC 430
 QY 488 AGGATTTCTTGTTCCTCAATTTGTTCCAAAGCGGTAACTTGCAGCAGGTAAATACA 547
 DB 431 AGGATTTCTTGTTCCTCAATTTGTTCCAAAGCGGTAACTTGCAGCAGGTAAATACA 490
 QY 548 CTGATGCTTAACCGTACCGTATCTACCAAGATTCAT-----GATTAGAGCCT 598
 DB 491 CTGATGCTTAACCGTACCGTATCTACCAAGATTCATCGAAGTTCGTTATAGACCT 550
 QY 599 ACGAACAACCAACCAACATTTTATGAGATCTAGAAAAAGTTAGGTGGAACAACCTTA 658
 DB 551 ACGAACAACCAACCAACATTTTATGAGATCTAGAAAAAGTTAGGTGGAACAACCTTA 610
 QY 659 CTGATGCTTAACCGTACCGTATCTACCAAGATTTGTTGAGTATGCAAGTATGGAAGTATG 718
 DB 611 CTGATGCTTAACCGTACCGTATCTACCAAGATTTGTTGAGTATGCAAGTATGGAAGTATG 670
 QY 719 AATTTTCAATTAATATGATCCAGAAAGATTCGAGGTTTGGCAATAGATTA 778
 DB 671 AATTTTCAATTAATATGATCCAGAAAGATTCGAGGTTTGGCAATAGATTA 730
 QY 779 CTGATGCTTAACCGTACCGTATCTACCAAGATTTTCTACCCGAGATACCATTC 838
 DB 731 CTGATGCTTAACCGTACCGTATCTACCAAGATTTTCTACCCGAGATACCATTC 790
 QY 839 TTAAGGCGGTCTTATGCAACCACTGCAAAATGCAATGCAAGTAAAGATTC 898

DB 791 TTAAGGCGGTCTTATGCAACCACTGCAAAATGCGATCAAGCGAGTAAAGATTC 850
 QY 899 TTAAGGCGGTCTTATGCAACCACTGCAAAATGCGAGTCAAGCGAGTAAAGATTC 958
 DB 851 TTAAGGCGGTCTTATGCAACCACTGCAAAATGCGAGTCAAGCGAGTAAAGATTC 910
 QY 959 CTTAAGCGGTCTTATGCAACCACTGCAAAATGCGAGTCAAGCGAGTAAAGATTC 1018
 DB 911 CTTAAGCGGTCTTATGCAACCACTGCAAAATGCGAGTCAAGCGAGTAAAGATTC 970
 QY 1019 ATCATGCTATGCGGTACCAAGTTCGTTGAAGATTTAGCTTACCGCGCAATTA 1078
 DB 971 ATCATGCTATGCGGTACCAAGTTCGTTGAAGATTTAGCTTACCGCGCAATTA 1030
 QY 1079 AGATTTATCAGTTATTTCAAGCGCAATTAATAGCATCTGCTAGTATGCGCACATTA 1138
 DB 1031 AGATTTATCAGTTATTTCAAGCGCAATTAATAGCATCTGCTAGTATGCGCACATTA 1090
 QY 1139 AATCCATGATTAATCCATTAATCTCATGATTAATTAATGATTAATGATTAAG 1198
 DB 1091 AATCCATGATTAATCCATTAATCTCATGATTAATTAATGATTAATGATTAAG 1150
 QY 1199 AGATTTTAAAGCGCGAGAGTACAAATTTCTGAGAAAAATGCTCAACCCCATTC 1258
 DB 1151 AGATTTTAAAGCGCGAGAGTACAAATTTCTGAGAAAAATGCTCAACCCCATTC 1210
 QY 1259 AGTGTATGAGAGCGAGAAAAATAGTCTCATTAAGAGCTTCTTGAAGTGAATA 1318
 DB 1211 AGTGTATGAGAGCGAGAAAAATAGTCTCATTAAGAGCTTCTTGAAGTGAATA 1270
 QY 1319 AAGAAACGGGGCGTGGTAAATCTGAATACTCATCTTATATAAAGTAAATAG 1378
 DB 1271 AAGAAACGGGGCGTGGTAAATCTGAATACTCATCTTATATAAAGTAAATAG 1330
 QY 1379 AATATCTACTTTGGCACCACTGCTCGGATTAAGTCCAGCGCGCTCAACGACTTGTTA 1438
 DB 1331 AATATCTACTTTGGCACCACTGCTCGGATTAAGTCCAGCGCGCTCAACGACTTGTTA 1390
 QY 1439 GCCAAAAACAACCTAGCTGCTGATATATCATCAAGTTTAACTAGTATGGAAGC 1498
 DB 1391 GCCAAAAACAACCTAGCTGCTGATATATCATCAAGTTTAACTAGTATGGAAGC 1450
 QY 1499 TGAACGTTTCAATGAAATATGATTCAGTATGATCAACGCTGCTAGATGACAGTCTG 1558
 DB 1451 TGAACGTTTCAATGAAATATGATTCAGTATGATCAACGCTGCTAGATGACAGTCTG 1510
 QY 1559 GTTAAATGA 1566
 DB 1511 GTTAAATGA 1518

RESULT 2
 US-08-913-477-16
 ; Sequence 16, Application US/08913477
 ; Patent No. 5985285
 ; GENERAL INFORMATION:
 ; APPLICANT: Titball, Richard W.
 ; APPLICANT: Williamson, Ethel D.
 ; APPLICANT: Leary, Sophie E.C.
 ; APPLICANT: Oyston, Petra C.F.
 ; APPLICANT: Bennett, Alice M.
 ; TITLE OF INVENTION: VACCINES FOR PLAGUE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHYE P.C.
 ; STREET: 1100 No. 5985285th Glebe Rd. 8th floor
 ; CITY: Arlington
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22201-4741
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1462 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 8..1447
US-08-913-477-16

```

```

Query Match          90.0%; Score 1409; DB 4; Length 1462;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1431; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

```

```

QY 131 CGGCGATTAACTGCAACGACCACTGCAACGGCAACTCTGTTGTAACGACCCGCATCA 190
DB 6 CGGCGATTAACTGCAACGACCACTGCAACGGCAACTCTGTTGTAACGACCCGCATCA 65
QY 191 CTTTACATATATAGGAAGGCGCTCCCAATTATGCAATTTAGCAATGGAACATCGATACAG 250
DB 66 CTTTACATATATAGGAAGGCGCTCCCAATTATGCAATTTAGCAATGGAACATCGATACAG 125
QY 251 AATTACTTGTGTAGCTTACTCTTGGCGGCTATAAAGAGAAACCACTAGACATCTG 310
DB 126 AATTACTTGTGTAGCTTACTCTTGGCGGCTATAAAGAGAAACCACTAGACATCTG 185
QY 311 TTTACTTTACAGATGCGGCGGCTGATCCATGTACTTAAACATTACTTCTCAGATGAA 370
DB 186 TTTACTTTACAGATGCGGCGGCTGATCCATGTACTTAAACATTACTTCTCAGATGAA 245
QY 371 ATTACCAACAATTCACTAATAAAGTATGGCAAGATTTCTAGATTTTGTATCTCTC 430
DB 246 ATTACCAACAATTCACTAATAAAGTATGGCAAGATTTCTAGATTTTGTATCTCTC 305
QY 431 CTAAGGTAACGCTGAGAACTTGTGGGGATGACGCTGCTGTGGCTAGCGGCAAGCAG 490
DB 306 CTAAGGTAACGCTGAGAACTTGTGGGGATGACGCTGCTGTGGCTAGCGGCAAGCAG 365
QY 491 ATTCTTTTGTGCTCAATTTGTTCCAAAGCGGTAAACTTGCAGCAGGTAAATACACTG 550
DB 366 ATTCTTTTGTGCTCAATTTGTTCCAAAGCGGTAAACTTGCAGCAGGTAAATACACTG 425

```

```

QY 551 ATGCTTAACGCTAACCCGATCTTAACCAAGAAATTCAT-----GATTAGAGCCTAGC 601
DB 426 ATGCTTAACGCTTAACCCGATCTTAACCAAGAAATTCATTCGAAGGTGTATTAAGCCTAGC 485
QY 602 AACAAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGGGTGAACAACACTACTG 661
DB 486 AACAAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGGGTGAACAACACTACTG 545
QY 662 GTCATGCTTCTTCAGTTTGAAGAATTGCTCAGTTAGTCAAGATTAATAAATATAGATA 721
DB 546 GTCATGCTTCTTCAGTTTGAAGAATTGCTCAGTTAGTCAAGATTAATAAATATAGATA 605
QY 722 TTTCCATTAAATATGATCCCAAAAAAGATGGAGGTTTTCCTCAATAGATAATACAG 781
DB 606 TTTCCATTAAATATGATCCCAAAAAAGATGGAGGTTTTCCTCAATAGATAATACAG 665
QY 782 ATGATATCGAATTGCTCAAGAAATCCTAGCTTATTTCTACCCGAGATACCATTTCTTA 841
DB 666 ATGATATCGAATTGCTCAAGAAATCCTAGCTTATTTCTACCCGAGATACCATTTCTTA 725
QY 842 AAGCGGCTATTATGACAACTGCAAAATGCAATGCAAGCAGATTAAGAAGTCTCTG 901
DB 726 AAGCGGCTATTATGACAACTGCAAAATGCAATGCAAGCAGATTAAGAAGTCTCTG 785
QY 902 AATCATCGCCGAATATACAAATGGGAATTGGGGCGCTCATGGCAGTAATGCATTTCTCT 961
DB 786 AATCATCGCCGAATATACAAATGGGAATTGGGGCGCTCATGGCAGTAATGCATTTCTCT 845
QY 962 TAACCGCCGATGATATGATGATATTTTGAAGTGAATGTTGATTAATGAATCATC 1021
DB 846 TAACCGCCGATGATATGATGATATTTTGAAGTGAATGTTGATTAATGAATCATC 905
QY 1022 ATGCTATGCCCGTAGCAAGTTCGTTGAAGAATTAGCTGAGCTTACCCGCAATTAAGA 1081
DB 906 ATGCTATGCCCGTAGCAAGTTCGTTGAAGAATTAGCTGAGCTTACCCGCAATTAAGA 965
QY 1082 TTTATTCATTTATTCAGCCGAAATTAATAGCATCTGCTAGTAGTGGCACCATTAATA 1141
DB 966 TTTATTCATTTATTCAGCCGAAATTAATAGCATCTGCTAGTAGTGGCACCATTAATA 1025
QY 1142 TCCATGATTAATTCATTAATTCATGATTAATAAATTTATATGTTATACAGATGAAGA 1201
DB 1026 TCCATGATTAATTCATTAATTCATGATTAATAAATTTATATGTTATACAGATGAAGA 1085
QY 1202 TTTTAAAGCCAGCGCAGATCAAAATTCGAGAAAAATGCTCAAAACCACTTCAG 1261
DB 1086 TTTTAAAGCCAGCGCAGATCAAAATTCGAGAAAAATGCTCAAAACCACTTCAG 1145
QY 1262 TGGATGGAGCGAGAAAAATATGCTGATTAAGGACTTTCTTGGAGTGAATTAATA 1321
DB 1146 TGGATGGAGCGAGAAAAATATGCTGATTAAGGACTTTCTTGGAGTGAATTAATA 1205
QY 1322 GAACGGGGCGTTGGTATCTGAATAAATCACTACCTATATATAAGTAAATATGAAT 1381
DB 1206 GAACGGGGCGTTGGTATCTGAATAAATCACTACCTATATATAAGTAAATATGAAT 1265
QY 1382 TATCTACTTTCACACACACTGCTGGATAGTCCAGGCGCTCAACGACTTGTAGGC 1441
DB 1266 TATCTACTTTCACACACACTGCTGGATAGTCCAGGCGCTCAACGACTTGTAGGC 1325
QY 1442 AAAAAAACAACCTAGCTGTCTGATATTAATCAACGTTTAAATTCAGCTATTGAAGCACTGA 1501
DB 1326 AAAAAAACAACCTAGCTGTCTGATATTAATCAACGTTTAAATTCAGCTATTGAAGCACTGA 1385
QY 1502 ACCGTTTCATGAGAAATATGATTCAGTATGCAACGCTGCTAGATGACAGCTCTGGA 1561
DB 1386 ACCGTTTCATGAGAAATATGATTCAGTATGCAACGCTGCTAGATGACAGCTCTGGA 1445
QY 1562 AATGA 1566
DB 1446 AATGA 1450

```

```

RESULT 3
US-08-913-477-1
: Sequence 1, Application US/08913477
: Patent No. 5985285
: GENERAL INFORMATION:
: APPLICANT: Tibbail, Richard W.
: APPLICANT: Williamson, Ethel D.
: APPLICANT: Leary, Sophie E.C.
: APPLICANT: Oyston, Petra C.F.
: APPLICANT: Bennett, Alice M.
: TITLE OF INVENTION: VACCINES FOR PLAGUE
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHAYE P.C.
: STREET: 1100 No. 5985285th Glebe Rd. 8th floor
: CITY: Arlington
: STATE: VA
: COUNTRY: USA
: ZIP: 22201-4741
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/913,477
: FILING DATE: 15-SEP-1997
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/GB96/00571
: FILING DATE: 13-MAR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9505059.7
: FILING DATE: 13-MAR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9518946.0
: FILING DATE: 15-SEP-1995
: APPLICATION DATA:
: APPLICATION NUMBER: GB 9524825.8
: FILING DATE: 05-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.
: REGISTRATION NUMBER: 25,327
: REFERENCE/DOCKET NUMBER: 124-599
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-816-4000
: TELEFAX: 703-816-4100
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1014 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Yersinia pestis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..987
: US-08-913-477-1

```

```

Query Match 62.48; Score 976.8; DB 4; Length 1014;
Best Local Similarity 99.88; Pred. No. 5.9e-276;
Matches 978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 587 TGATTAGAGCTTACGACAAACCCACAAACATTTATGAGAGCTGAAAAAGTTAGGG 646
DB 11 TCATTAGAGCTTACGACAAACCCACAAACATTTATGAGAGCTGAAAAAGTTAGGG 70

```

```

QY 647 TGAACAACCTACTGTCATGTTCTTCACTTTAGAGAATTGGTCACTAGTCAAG 706
DB 71 TGAACAACCTACTGTCATGTTCTTCACTTTAGAGAATTGGTCACTAGTCAAG 130
QY 707 ATAAAAATATGATATTTCCATTAATATATGATCCAGAAAAAGATCGAGAGTTTGGCA 766
DB 131 ATAAAAATATGATATTTCCATTAATATATGATCCAGAAAAAGATCGAGAGTTTGGCA 190
QY 767 ATAGATTAATTACTGATGATATGATATGATGCTCAAGAAATCTAGCTTATTTCTACCG 826
DB 191 ATAGATTAATTACTGATGATATGATATGATGCTCAAGAAATCTAGCTTATTTCTACCG 250
QY 827 AGATACCACTTTAAAGGGGCTATTATGACAACAACGCAAAATGCGATCAAGCGAG 886
DB 251 AGATACCACTTTAAAGGGGCTATTATGACAACAACGCAAAATGCGATCAAGCGAG 310
QY 887 TAAAGAGTCTCTGATTCATCGCCGAATACACAATGGGAATTGCGGCGTTATGCGAG 946
DB 311 TAAAGAGTCTCTGATTCATCGCCGAATACACAATGGGAATTGCGGCGTTATGCGAG 370
QY 947 TAATGATTTCTCTTTAACCGCGATCGTATCGATGATGATTTTGAAGTGATGTTG 1006
DB 371 TAATGATTTCTCTTTAACCGCGATCGTATCGATGATGATTTTGAAGTGATGTTG 430
QY 1007 ATTCATGAAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
DB 431 ATTCATGAAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 490
QY 1067 CCGCGCAATTAAGATTTATTCAGTTTATCAAGCCGAATTAATTAAGATCTGCTAGTA 1126
DB 491 CCGCGCAATTAAGATTTATTCAGTTTATCAAGCCGAATTAATTAAGATCTGCTAGTA 550
QY 1127 GTGCGACCATTAATATCATGATTAATCATTAATCATGATGATTAATTAATTAATGTT 1186
DB 551 GTGCGACCATTAATATCATGATTAATCATTAATCATGATGATTAATTAATTAATGTT 610
QY 1187 ATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTGAGAAATGCTC 1246
DB 611 ATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTGAGAAATGCTC 670
QY 1247 AAACACCATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1306
DB 671 AAACACCATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 730
QY 1307 GAATGAGATTAATTAAGAACCGGCGGTTGGTATCTGAATAAATCTACTACTTATATA 1366
DB 731 GAATGAGATTAATTAAGAACCGGCGGTTGGTATCTGAATAAATCTACTACTTATATA 790
QY 1367 AAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1426
DB 791 AAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 850
QY 1427 ACGACTGTGTTAGCCAAACAACTAGCTGCTGATTAATTAATTAATTAATTAATTAAT 1486
DB 851 ACGACTGTGTTAGCCAAACAACTAGCTGCTGATTAATTAATTAATTAATTAATTAAT 910
QY 1487 CTATTGAAGCACTGACCGTTTCATTCAGAAATATGATGATGATGATGATGATGATGAT 1546
DB 911 CTATTGAAGCACTGACCGTTTCATTCAGAAATATGATGATGATGATGATGATGATGAT 970
QY 1547 ATGACACGCTGTGTAATGA 1566
DB 971 ATGACACGCTGTGTAATGA 990

```

```

RESULT 4
US-08-913-477-3
: Sequence 3, Application US/08913477
: Patent No. 5985285
: GENERAL INFORMATION:
: APPLICANT: Tibbail, Richard W.
: APPLICANT: Williamson, Ethel D.
: APPLICANT: Leary, Sophie E.C.

```

APPLICANT: Oyston, Petra C.F.
 APPLICANT: Bennett, Alice M.
 TITLE OF INVENTION: VACCINES FOR PLAGUE
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHAYE P.C.
 STREET: 1100 NO. 5985285th Gleda Rd. 8th floor
 CITY: Arlington
 STATE: VA
 COUNTRY: USA
 ZIP: 22201-4741
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/913,477
 FILING DATE: 15-SEP-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/GB96/00571
 FILING DATE: 13-MAR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9505059.7
 FILING DATE: 13-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9518946.0
 FILING DATE: 15-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9524825.8
 FILING DATE: 05-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Crawford, Arthur R.
 REGISTRATION NUMBER: 25,327
 REFERENCE/DOCKET NUMBER: 124-599
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-816-4100
 TELEFAX: 703-816-4100
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1014 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Yersinia pestis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..987
 OS-08-913-477-3

	Query Match	62.2%	Score 974.8	DB 4	Length 1014
	Best Local Similarity	99.8%	Pred. No. 2.3e-275		
	Matches 976	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	589	ATTAGAGCCTACGACAAAAACCACACATTTTATTGAGACTAGAAAAAGTAGGAGTG	648		
Db	13	ATTGAGCCTACGACAAAAACCACACATTTTATTGAGAGTCTAATAAAAGTTAGGAGTG	72		
QY	649	GACCAACTACGCGCATAGGTCTTCACGTTTAGAAGATTGTTAGTACCAAGAT	708		
Db	73	GACCAACTACGCGCATAGGTCTTCACGTTTAGAAGATTGTTAGTACCAAGAT	132		
QY	709	AAAAATATAGATATTTCCTTAATAATATGATCCAGAAAAGATCCGAGGTTTGGCAAT	768		
Db	133	AAAAATATAGATATTTCCTTAATAATATGATCCAGAAAAGATCCGAGGTTTGGCAAT	192		
QY	769	AGAGTAATTAAGTATGATATGCAATTTGCTCAAGAAAAATCCTACGCTTAATTTTACCCGAG	828		

Db	193	AGAGTAATTA	CTGATGAT	CGAATTG	CTCAAGAAAAT	CTTGCTTA	TTTCTTACC	GAG	252			
Qy	829	GATACCA	TTCTTAAAG	CGGCTCAT	TATGACAA	CCA	CTGCCAAAAT	TGCATCA	GGCAGTA	888		
Db	253	GATCCCA	TTCTTAAAG	CGGCTCAT	TTATGACA	CA	CCA	CTGCCAAAAT	TGCATCA	GGCAGTA	312	
Qy	889	AAAAGAT	CTCTGAA	TATATG	CGCGAAT	TACACAT	TGGGAAT	TGCGGGCT	TATG	CAGTA	948	
Db	313	AAAAGAT	CTCTGAA	TATATG	CGCGAAT	TACACAT	TGGGAAT	TGCGGGCT	TATG	CAGTA	372	
Qy	949	ATGCAT	TTCTCTTAA	CCGCCGAT	CGTATCGAT	GATGAT	TATTTTGA	AAATGAT	TTTGT	TGAT	1008	
Db	373	ATGCAT	TTCTCTTAA	CCGCCGAT	CGTATCGAT	GATGAT	TATTTTGA	AAATGAT	TTTGT	TGAT	432	
Qy	1009	TCAATGA	TATCAT	CAATG	TGTGAT	GCCCGTAG	CA	AGTTGCT	GAAGAA	TTTAC	TGACGCTTAC	1068
Db	433	TCATATGA	TATCAT	CAATG	TGTGAT	GCCCGTAG	CA	AGTTGCT	GAAGAA	TTTAC	TGACGCTTAC	492
Qy	1069	GCCCAAT	TTAAAGAT	TTTATTA	GTATTTCA	AGCCGGA	ATTAAT	TATAG	CATCTG	CTGTAGT	1122	
Db	493	GCCCAAT	TTAAAGAT	TTTATTA	GTATTTCA	AGCCGGA	ATTAAT	TATAG	CATCTG	CTGTAGT	552	
Qy	1129	GGCACCA	TAATAT	TCATGAT	CGATTAAT	CCATTAAT	CTCATG	ATGATTA	AAATTTAT	TATG	TAT	1188
Db	553	GGCACCA	TAATAT	TCATGAT	CGATTAAT	CCATTAAT	CTCATG	ATGATTA	AAATTTAT	TATG	TAT	612
Qy	1189	ACAGATGA	AGAGAT	TTTTTAA	AGCCAGC	AGATAC	CAAAAT	TTCTG	AGAAAAT	TGCTTAA	1248	
Db	613	ACAGATGA	AGAGAT	TTTTTAA	AGCCAGC	AGATAC	CAAAAT	TTCTG	AGAAAAT	TGCTTAA	672	
Qy	1249	ACCACCA	TTCAAGT	GTGATG	GAGCGA	GGAAAAAAT	ATGTCG	ATTAAG	CAATCTT	CTTGA	1308	
Db	673	ACCACCA	TTCAAGT	GTGATG	GAGCGA	GGAAAAAAT	ATGTCG	ATTAAG	CAATCTT	CTTGA	732	
Qy	1309	AGTAGAAT	TAAGAA	ACC	GGGGCGT	TGGTAT	CTGAAA	AACTCAT	TCTTAT	ATAA	1368	
Db	733	AGTAGAAT	TAAGAA	ACC	GGGGCGT	TGGTAT	CTGAAA	AACTCAT	TCTTAT	ATAA	792	
Qy	1369	GATATAT	TATGAT	TATAT	CTACTT	TGACAC	ACACTG	TGCGAT	TAAGCA	GGCGGCT	CAC	1428
Db	793	GATATAT	TATGAT	TATAT	CTACTT	TGACAC	ACACTG	TGCGAT	TAAGCA	GGCGGCT	CAC	852
Qy	1429	GACTTGGT	TATAGCC	AAAAA	AACTCA	AGCTGT	CGATTA	TACATCA	CGTTTAA	TACAGT	1488	
Db	853	GACTTGGT	TATAGCC	AAAAA	AACTCA	AGCTGT	CGATTA	TACATCA	CGTTTAA	TACAGT	912	
Qy	1489	ATTGAGAC	ACTGAAC	CGTTTCAT	TACAGAA	TATATG	ATTCAGT	AGTAC	GAACAGT	CTGCTAGAT	1548	
Db	913	ATTGAGAC	ACTGAAC	CGTTTCAT	TACAGAA	TATATG	ATTCAGT	AGTAC	GAACAGT	CTGCTAGAT	972	
Qy	1549	GACACG	CTCTG	TAAATGA	1566							
Db	973	GACACG	CTCTG	TAAATGA	990							

RESULT 5
US-08-913-477-20
Sequence 20, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 No. 5985285th Glede Rd.
CITY: Arlington
STATE: VA

RESULT 13

US-08-308-872B-5
Sequence 5, Application US/08308872B
Patent No. 5661006

GENERAL INFORMATION:

APPLICANT: BROWN, Thomas David Kay
APPLICANT: HORSBURGH, Brian Collin
TITLE OF INVENTION: CANINE CORONAVIRUS SUBUNIT VACCINE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Akzo No. 5661006el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,872B

FILING DATE: 19-SEP-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,846

FILING DATE: 05-APR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,641

FILING DATE: 24-APR-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91.303.737.0

FILING DATE: 25-APR-1991

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gormley, Mary E.

REGISTRATION NUMBER: 34,409

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 258-5200

INFORMATION FOR SEQ. ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 4435 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Canine corona virus

STRAIN: CCV-V54

FEATURE:

NAME/KEY: CDS

LOCATION: 60..4418

OTHER INFORMATION: /label=CCV-c54_Spikegene

US-08-308-872B-5

Query Match 2.5%; Score 39; DB 1; Length 4435;

Best Local Similarity 48.8%; Pred. No. 0.12; Indels 0; Gaps 0;

Matches 105; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 622 ATTGAGATCTAGAAAAGTTAGGTTGGAACAATTACTGATGCTTTCAGTTTGA 681
DB 120 AATAAAGATGTGACAAAGTTAATGTACACAAATGCTGCAATGAAATATCATTTAA 179
QY 682 GAAGAATGTGTAGTTCAGTCAAGATTAATAATATAGATTTCCATTAATATGATCCC 741
DB 180 GATTTCTATATTTAGATTTTAAAGAGAGAGTGTAGTTGTGCTACTACCCC 239
QY 742 AGAAAGATTCGGAGGTTTGGCAATAGAGTATTTCTGATGATGATGCAATGCTCAAG 801
DB 240 ACAGAGGTGTGTAACAAGTGTTCAGACAGACACAACTACAGCTTACCATTTATTAGT 299

RESULT 14

US-08-945-056-1
Sequence 1, Application US/08945056
Patent No. 6077994

GENERAL INFORMATION:

APPLICANT: Coupland, George M.

APPLICANT: Putterill, Joanna J.

TITLE OF INVENTION: Genetic control of flowering

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSER: Nixon & Vanderyhe PC

STREET: 8th Floor, 1100 No. 6077994th Glebe Road

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,056

FILING DATE: 20-OCT-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02561

FILING DATE: 01-NOV-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9422083.7

FILING DATE: 02-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Ms Mary J Wilson

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 620-17

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1124 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Arabidopsis thaliana

STRAIN: Landsberg erecta

POSITION IN GENOME:

CHROMOSOME/SEGMENT: chromosome 5

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1119

US-08-945-056-1

Query Match 2.5%; Score 38.6; DB 5; Length 1124;

Best Local Similarity 52.1%; Pred. No. 0.075; Indels 0; Gaps 0;

Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1 ATGGGCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
DB 664 AGGGGCCACACAGTCATTAACAGAAATTTTCAGTTCATTAACAAATATGCTCCTCA 723
QY 61 GACAACATATGAAAAAATCAATTCGCTTATTCGCAATTCATTTATTTGGAATATGCA 120
DB 724 GAGCACTACTACACAGCAATGTTCCATTAACCAATACGCAATACATTCATCATGGA 783

QY 121 ACTGCTAATGCGGACAGATTACTGCAAGCACCAGTCGACGCA 165
| | | | | | | | | | | | | | | | | | | | | |
Db 784 ACTGCTGTTGTGCCGAGTCAACAGCATGTGTGACACAGCTTCA 828

RESULT 15

PCT-US91-08525-31
: Sequence 31, Application PC/TUS9108525
: GENERAL INFORMATION:
: APPLICANT: SmithKline Beecham, Corporation
: TITLE OF INVENTION: Recombinant feline Coronavirus S
: NUMBER OF SEQUENCES: 54
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406-2799
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US91/08525
: FILING DATE: 19911114
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/698,927
: FILING DATE: 13-MAY-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/613,066
: FILING DATE: 14-NOV-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: King, William T.
: REGISTRATION NUMBER: 30,954
: REFERENCE/DOCKET NUMBER: SRC 14532B
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 270-5015
: TELEFAX: (215) 270-5090
: INFORMATION FOR SEQ ID NO: 31:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4365 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: unknown
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..4362
: PCT-US91-08525-31

Query Match

2.48: Score 37.4; DB 6; Length 4365;
Best Local Similarity 47.68; Pred. No. 0.35; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 121;

QY 626 AGATCTAGTAAAGTGTAGGAGCAACTACTGTCAGTTTGAAG 685
| | | | | | | | | | | | | | | | | | | | | |
Db 68 ATGATGTAGACAGTACGTAACACAAATGCTGCAATGAAACCTTATAGAGACT 127
QY 686 AATGGTCACTTACTCAAGATAAAATATAGATATTTCCATTAAATATAGATCCAGAA 745
| | | | | | | | | | | | | | | | | | | | | |
Db 128 TTTGTTCAAGTTTAAAGAGAAAGATGTGTGTTGTTGTTATTAACCTACAG 187
QY 746 AAGATTCGAGGTTTTCGCAATAGATATGATGATTCGATTCGATTCGCAAGAAA 805
| | | | | | | | | | | | | | | | | | | | | |
Db 188 AGGTGTGTAACACTGCTTGAACAGCACTACCTGCTATGATTTTAAATAATA 247
QY 806 TCCTAGCTTATTTCTACCCAGAGATACCATCTTAAAGCGGTATATG 856

Db 248 TACATGCCCTTTATTTGATATGAAAGCTATGAAATAGCAGCTGTATG 298
| | | | | | | | | | | | | | | | | | | | | |

Search completed: August 22, 2000, 15:11:41
Job time: 780 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 22, 2000, 14:56:42 ; Search time 613.89 Seconds
(without alignments)
11249.232 Million cell updates/sec

Title: US-08-699-716a-1

Perfect score: 1566
Sequence: 1 ATGGGCCATCATCATCA.....ATGACAGCTCTGTAAATGA 1566

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_est56:*
94: gb_est57:*
95: gb_est58:*
96: gb_est59:*
97: em_est60:*
98: em_est61:*
99: em_est62:*
100: em_est63:*
101: gb_est64:*
102: gb_est65:*
103: gb_est66:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: em_est60:*
108: em_est61:*
109: em_est62:*
110: em_est63:*
111: em_est64:*
112: em_est65:*
113: gb_est70:*
114: gb_est71:*
115: em_est72:*
116: gb_est73:*

```

117: qb_gss13:*
118: qb_gss14:*
119: qb_gss15:*
120: qb_gss16:*
121: qb_gss17:*
122: qb_gss18:*
123: qb_gss19:*
124: em_gss13:*

```

Description

C	1	50.6	3.2	1101	122	CNS00E2P	AL069652	Drosophila
C	2	48.8	3.1	1101	122	CNS0039G	AL068391	Drosophila
C	3	48	3.1	1201	122	CNS0102D	AL068431	Drosophila
C	4	47	3.0	1101	122	CNS0038G	AL063921	Drosophila
C	5	44	2.8	797	122	CNS00038	AL064634	Drosophila
C	6	42.6	2.7	700	117	AQ940248	AQ940248	Sheared D
C	7	41.8	2.7	579	48	AU038989	AU038989	AU038989
C	8	41.8	2.7	1101	123	CNS0182P	AL108811	Drosophila
C	9	41.6	2.7	797	122	CNS0038W	AL064634	Drosophila
C	10	41.6	2.7	1101	122	CNS0143H	AL103598	Drosophila
C	11	41.4	2.6	791	96	AQ258723	AQ258723	npxb0021G
C	12	41	2.6	444	86	H78200	H78200	yu84f10.r1
C	13	41	2.6	964	122	CNS0069	AL065781	Drosophila
C	14	41	2.6	1101	122	CNS00LVP	AL078809	Drosophila
C	15	40.8	2.6	561	81	C91275	C91275	C91275
C	16	40.8	2.6	1101	122	CNS0006J	AL062049	Drosophila
C	17	40.6	2.6	394	87	N97614	N97614	1054C3 czap
C	18	40.6	2.6	861	122	CNS0116N	AL069881	Drosophila
C	19	40.2	2.6	1201	123	CNS0165X	AL106335	Drosophila
C	20	40	2.6	890	123	CNS0178S	AL108346	Drosophila
C	21	40	2.6	1101	123	CNS0178S	AL107617	Drosophila
C	22	39.8	2.5	372	116	AQ687381	AQ687381	nbebu031D
C	23	39.8	2.5	544	116	AQ626892	HS-5265_B	AQ626892
C	24	39.4	2.5	534	118	AQ982815	AQ982815	RPCI-23-B
C	25	39.2	2.5	395	93	AQ037655	AQ037655	Cir-HSP-2
C	26	39.2	2.5	827	122	CNS00E3	AL069854	Drosophila
C	27	39.2	2.5	1200	123	CNS0166O	AL106578	Drosophila
C	28	39	2.5	510	81	C84881	C84881	C84881
C	29	39	2.5	889	122	CNS00D9	AL067197	Drosophila
C	30	38.6	2.5	1000	122	CNS005CQ	AL059446	Drosophila
C	31	38.6	2.5	273	73	AM505157	AM505157	UI-HF-BNO
C	32	38.6	2.5	489	47	AL121551	AL121551	DRE2P762L
C	33	38.6	2.5	526	62	AM004525	AM004525	70193208
C	34	38.6	2.5	924	103	AQ050007	AQ050007	HS-5202_B
C	35	38.6	2.5	625	122	CNS0056T	AL057103	Drosophila
C	36	38.6	2.5	1086	122	CNS00YXK	AL086962	Drosophila
C	37	38.6	2.5	1107	80	C23232	C23232	C23232
C	38	38.6	2.5	1203	123	CNS015WU	AL106008	Drosophila
C	39	38.4	2.5	821	114	AQ795386	AQ795386	nxbx0056B
C	40	38.2	2.4	560	113	AQ727757	AQ727757	HS-5457_B
C	41	38.2	2.4	769	103	AQ487444	AQ487444	RRC1-11-2
C	42	38.2	2.4	1086	122	CNS00YXK	AL086962	Drosophila
C	43	38.2	2.4	1101	123	CNS017XK	AL108711	Drosophila
C	44	38	2.4	243	28	AA547782	AA547782	EST188691
C	45	38	2.4	522	113	AQ680721	AQ680721	HS-3481_B

DEFINITION	LOCUS	RESULT 1
BACR30E19 of RPci-98 library from <i>Drosophila melanogaster</i> (Truh	1101 bp	04-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC:	DNA	GSS
CNS00E2P		
CNS00E2P/c		

COMMENT

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information, please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Oseegawa and Aaron Mammosier in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCR-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

ORGANISM

LOCUS	CNS000396	1101 bp	DNA	GSS	03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR88k10 of RPC1-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.				
ACCESSION	AL063921				
VERSION	AL063921.1				
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				

CNS00396/c
LOCUS 1101 bp DNA GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063921 GI:4941778
VERSION AL063921.1
KEYWORDS GSS.
ORGANISM fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL - Web : www.genoscope.cns.fr
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila melanogaster](http://www.fruitfly.org/TheBDGP/Drosophila%20melanogaster%20BAC%20library%20was%20prepared%20by%20Kazutoyo%20Osoegawa%20and%20Aaron%20Mammoser%20in%20Pleier%20de%20Jong's%20laboratory%20in%20the%20Department%20of%20Cancer%20Genetics%20at%20the%20Roswell%20Park%20Cancer%20Institute%20in%20Buffalo,%20NY.%20The%20library%20is%20named%20RPCI-98%20and%20was%20constructed%20by%20partial%20EcoRI%20digestion%20of%20Drosophila%20DNA%20provided%20by%20the%20BDGP%20from%20the%20isogenic%20strain%20y2.%20cn%20bw%20sp.%20The%20same%20strain%20used%20for%20the%20BDGP's%20P1%20and%20EST%20libraries.%20A%20more%20detailed%20description%20of%20the%20library%20and%20how%20to%20order%20individual%20BAC%20clones,%20the%20entire%20library,%20or%20filters%20for%20hybridization%20from%20the%20BACPAC%20Resource%20Center%20can%20be%20found%20at%20http://bacpac.med.buffalo.edu/drosophila_bac.htm)
FEATURES
SOURCE 1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_1id="RPCI-98"
/clone="BACR08K10"
/note="end : TET3"
BASE COUNT 201 a 64 c 131 g 202 t 503 others
ORIGIN

Query Match 3.0%; Score 47; DB 122; Length 1101;
Best Local Similarity 14.1%; Pred. No. 0.057;
Matches 59; Conservative 198; Mismatches 162; Indels 0; Gaps 0;

OY 974 GTATGATGATGATATTTTGAAGTATGATTCATCATGATCATGATGATGCC 1033
DB 946 KWTATWMDWDMDKMDKMDKMDKMDKMDKMDKMDKMDKMDKMDKMDKMD 887
OY 1034 GTAGAGTGTGGTGAAGATAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTA 1093
DB 886 DDMWDRAKGTGAT 827
OY 1094 TTCAAGCCGAATTAATAGATCTGTCTAGTGTAGTGTAGTGTAGTGTAGTGT 1153
DB 826 WGAFTADRDWDGDRGKRGKRGKRGKRGKRGKRGKRGKRGKRGKRGKRGK 767
OY 1154 CCATTAACTCTGATATAAATTTATATATGTTATACAGATAGAGATTTTAAAGCA 1213
DB 766 KWKLTWTWADRTWADDDDDDRAGATGKRRKRRKRRKRRKRRKRRKRRKRR 707
OY 1214 GCGCAGAGACAAATTTCTCGAAGAAATCCCAAAACCAATTCAGCGATGGAGAG 1273
DB 706 DRRRGDGDGADGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 647
OY 1274 AGAAAAAATAGCTGATTAAGACTTTCTTGAAGTGAAGTGAAGTGAAGTGAAG 1333
DB 646 KGARRRRRTTAAGAAADWMTWKMDAKMDKMDKMDKMDKMDKMDKMDKMDKMD 587

OY 1334 TGGTATCTGAATACTGATCTCTATATATAAGATATATGAAATATCTGACTT 1392
DB 586 WRARDRAR 528

RESULT 5
LOCUS 797 bp DNA GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08C03 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL064634 GI:4941986
VERSION AL064634.1
KEYWORDS GSS.
ORGANISM fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 797)
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL - Web : www.genoscope.cns.fr
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila melanogaster](http://www.fruitfly.org/TheBDGP/Drosophila%20melanogaster%20BAC%20library%20was%20prepared%20by%20Kazutoyo%20Osoegawa%20and%20Aaron%20Mammoser%20in%20Pleier%20de%20Jong's%20laboratory%20in%20the%20Department%20of%20Cancer%20Genetics%20at%20the%20Roswell%20Park%20Cancer%20Institute%20in%20Buffalo,%20NY.%20The%20library%20is%20named%20RPCI-98%20and%20was%20constructed%20by%20partial%20EcoRI%20digestion%20of%20Drosophila%20DNA%20provided%20by%20the%20BDGP%20from%20the%20isogenic%20strain%20y2.%20cn%20bw%20sp.%20The%20same%20strain%20used%20for%20the%20BDGP's%20P1%20and%20EST%20libraries.%20A%20more%20detailed%20description%20of%20the%20library%20and%20how%20to%20order%20individual%20BAC%20clones,%20the%20entire%20library,%20or%20filters%20for%20hybridization%20from%20the%20BACPAC%20Resource%20Center%20can%20be%20found%20at%20http://bacpac.med.buffalo.edu/drosophila_bac.htm)
FEATURES
SOURCE 1..797
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_1id="RPCI-98"
/clone="BACR08C03"
/note="end : TET3"
BASE COUNT 170 a 78 c 89 g 133 t 327 others
ORIGIN

Query Match 2.8%; Score 44; DB 122; Length 797;
Best Local Similarity 13.6%; Pred. No. 0.33;
Matches 54; Conservative 176; Mismatches 167; Indels 0; Gaps 0;

OY 603 ACAAAACCCACAACATTTTATGAGATCTAGAAAAGTGGTGAACAACATCTAG 662
DB 398 WBMAAANVATSSVMSATSHHWVDTWTWKKAHVHKWKMDKKKKVMMMAADRSK 457
OY 663 TCATGCTCTCAGTTTGAAGAATTTGTTCACTAGTCAAGATATAAATATAGATAT 722
DB 458 NAKKMDAAAAAAAVKMDKRRKKKMDKRRKKKMDKRRKKKMDKRRKKKMDKRR 517
OY 723 TTCCATTAATATGATCCCAAGAAATTCGAGATTTTCCATAGAGTAATAGTGA 782
DB 518 NNNNNNNKAAATATKSAKKDMSNATBHTKDKAKKBNKCBRAADKXVSAVAAAKK 577
OY 783 TGATATCAATTCGCAAGAAATCTAGCTTATTTCTACCGAGATACCATCTTAA 842
DB 578 GGGKRAVYNNVDVBSMAATRAMGKSVAAAMKDRSSNTBBDJTDVHVHYDMAAA 637
OY 843 AGCGGCTATATGCAACCACTGCAAAATGCGATCAAGGAGTAAAGAGTTCCTTGA 902

```

Db      638  DNYMAMVYKKAAMVYMAAAMMAAAMMAAMKDMAAAMAMDMAAMMWTBKKBDHKKRAMOM 697
OY      903  ATCATCGCCGATATACCAATGCGAATTCGGCGGCGTTCATGCAGTAATCAATTCCTTT 962
Db      698  WMMCCMMAAAMAAAMWMAKSCMKTRADPCMAKTRVCMCHTKDDMAAAMCMWYAYB 757
OY      963  AACCCGCGATCGATGATGATGATATATTTGAAGTG 999
Db      758  MMCCCYAAYKNTWNTNHTDMDMTRKMAAMAKND 794

RESULT      6
AC0940248
LOCUS
DEFINITION  Sheared DNA-4ZE21.TF Sheared DNA Trypanosoma brucei genomic
ACCESSION  clone sheared DNA-4ZE21, genomic survey sequence.
              AC0940248
VERSION      AC0940248.1 GI:6763513
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
ORGANISM    Trypanosoma brucei.
REFERENCE   1 (bases 1 to 700)
AUTHORS     El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Sub,E., Malek,J., Fujita,C.,
            Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S.,
            Donelson,J., Fraser,C. and Adams,M.
            Determination of clone end sequences from Trypanosoma brucei GUTat
            10.1 sheared DNA library
            Unpublished (1999)
JOURNAL
COMMENT      Other_GSSs: Sheared DNA-4ZE21.TR
            Contact: Najib M. El-Sayed
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: nelsayed@tigr.org
            Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
            DNA library constructed at TIGR. Clones will be available for
            distribution through ATCC. Sheared DNA end sequences search page:
            http://www.tigr.org/tdb/mdb/tbdb/.
            Seq primer: M13-Forward
            Class: Shotgun.
FEATURES
            source
            Location/Qualifiers
                1..700
                /organism="Trypanosoma brucei"
                /strain="TREU927/4 GUTat 10.1"
                /db_xref="taxon:5691"
                /clone="Sheared DNA-4ZE21"
                /clone_1lb="Sheared DNA"
                /note="Vector: pUC18. Site_1: SmaI. Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
                Genomic DNA isolated from a cloned population of
                Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
                sheared to give a tight size distribution (approx 2 kb).
                The v + i method used for the library construction is
                described in detail in Smith, H.O. and Venter, J.C.
                (Making small insert libraries for whole genome shotgun
                sequencing projects. In Genome Sequencing: A Practical
                Approach, eds. M. Vaundin and B. Barrell, Oxford University
                Press, 1999)."
```

DB	220	TCGACACTGTATTAATTAAGTGGCACACATTAAGAATGTTATGTTAATAGACATTAATA	279
QY	975	TATGCATGATGATTAATTTGAAAAGTATGTTGATTCATCAATGAATCATCATGATGCGCCG	1034
DB	280	TAAAT	339
QY	1035	TAGCAAGTTGGCTGAGAGATTAGCTGAGCTTACCCCGAATTAAAGATTATATTCAGTTAT	1094
DB	340	TAAAT	399
QY	1095	TCNAGCCGAAATTAATTAAGCATCTGCTAGTAGTGCGACCATTAATATTCATGATTAATTC	1154
DB	400	TAAAT	459
QY	1155	CATTAACTCATGATTAATAAATTTATATGTTATACAGATGAGAGATTTTAAAGCCAG	1214
DB	460	TAAAT	519
QY	1215	CGCAGAGTACAAATATTCGAGAAAATGCGCTCAAAACCACCATTCAGGTCGATGGACCA	1274
DB	520	TAAAT	579
QY	1275	GAAAAAATAGTCTCGATGAAGAGCTTCTTGGAAGTGAAGTAAGAAGACGGGGCGTT	1334
DB	580	TAAAT	639
QY	1335	GGGTAATCTGAAAAACCTCATCTCTTAATTAATTAAGATATATATGA	1379
DB	640	TAAAT	684

RESULT	7
LOCUS	AU038989
DEFINITION	AU038989 Dictyostelium discoidium SS (H. Urushihara) Dictyostelium
ACCESSION	AU038989
VERSION	AU038989.1
KEYWORDS	GI:3985742
SOURCE	EST.
ORGANISM	Dictyostelium discoidium.
REFERENCE	Dictyostelium discoidium.
AUTHORS	Eukaryota; Dictyostellida; Dictyostelium.
	1 (bases 1 to 579)
	Morita, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,
	Yoshino, R., Mita, B. N., Pl. M., Saito, T., Takemoto, K., Yasukawa, H.,
	Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
	The Dictyostelium developmental cDNA project: generation and
	analysis of expressed sequence tags from the first-finger stage of
	development
	DNA Res. 5 (6), 335-340 (1998)
JOURNAL	
MEDLINE	99156227
COMMENT	On Jun 5, 1998 this sequence version replaced gi:3189400.
	Contact: Hideko Urushihara
	Institute of Biological Sciences
	University of Tsukuba
	3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan
	Email: d402huesakura.cc.tsukuba.ac.jp
	PROJECT - 'Dictyostelium discoidium cDNA project in Japan'.
FEATURES	
SOURCE	Location/Qualifiers
	1..579
	/organism="Dictyostelium discoidium"
	/strain="AX4"
	/db_xref="taxon:44689"
	/clone="SSM246"
	/clone_1lb="dictyostelium discoidium SS (H. Urushihara)"
	/dev_stage="slug"
BASE COUNT	286 a 31 c 52 g 210 t
ORIGIN	

Query Match	2.7%	Score 41.8:	DB 48:	Length 579:
Best Local Similarity	45.0%	Pred. No. 1.1:		
Matches 157: Conservative	0:	Mismatches 192:	Indels 0:	Gaps 0:

QY	970	GATCGTATTCAGTAATGATATTTTGAAAGCATGGTGTGATTCAATCAGTATCGTGTAT	1029
Db	71	GCTCATATTTTTTGATGACAAAATTTAATAATTTTCATAGTAATAATAATAATAATGCAT	130
QY	1030	GCCCGTAGCAAGTTGCGTGAGAATTTAGCTGAGCTTACCGCCGAATTTAAGATTATCCA	1089
Db	131	GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	190
QY	1090	GTTATTCAGCCGAATTTAATAGATCTGCTAGTAGAGGCACATATAATTCATGAT	1149
Db	191	AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	250
QY	1150	AAATCCATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1209
Db	251	AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	310
QY	1210	GCCAGCGCAGAGTACAAAAATTCGAGAAAATGCTCAAACACCATTCAGGTGATGGG	1269
Db	311	TGGATTAAACAATGTAAATGATATGATCAATCAATATTTTCAATTAATAATCAAGAGTGT	370
QY	1270	AGCGAGAAAAAATAGTCTCGATTAAGAGCTTTCTTGAGAGGAGAAATA	1318
Db	371	ATGAACAAGAGTGTGTTATTTAAAAATGAATTAATAATGAATTAATGAATA	419
RESULT	8		
LOCUS	CNS0182P	1101 bp	DNA
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BAON37D10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL108811		
VERSION	ALI08811.1	GI:5629115	
KEYWORDS	GSS.		
SOURCE	GSS.		
ORGANISM	Fruit fly. Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope. Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CBPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.		
COMMENT			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source	Location/Qualifiers		
	1..1101		
	/organism="Drosophila melanogaster"		
	/plasmid="pBelobAC11"		
	/db_xref="taxon:7227"		
	/clone_11b="DrosBAC"		
	/clone="BAON37D10"		
	/note="end : SP6"		
BASE COUNT	274 a 268 c 128 g 73 t 358 others		
ORIGIN			
	Query Match 2.7%; Score 41.8; DB 123; Length 1101;		
	Best Local Similarity 17.1%; Pred. No. 1.4;		
	Matches 62; Conservative 164; Mismatches 133; Indels 4; Gaps 1.		
QY	979	GATGATGATATTTTAAAGTGTGTTGATTTAAAGATCAATCAGTATCCCCGTAGC	1038
Db	742	GKKDWGTAAWMTWATWADTWKATTDIDAKRAAAGRRRDARRKARDGRRAATRRRAW	801

QY	1039	AAGTGGCGTGAAGATTGCGAGCTTGCAGCCCGCATTAAGATTATCGATTATCA	1098
Db	802	AAGGRRARAAGARRARAARRADDNDNDMAAAAAAAAAAAATTTNDNRWMDNDMDNR	861
QY	1099	GCCGAATTAATAAGCATCTGCTAGTAGTGGCACCATTAATATCATGATTAATCCATT	1158
Db	862	D-----DTTTTAAWMDPARARARRRRRRRRRRARRARAARRADDNDTKRWADATTDKDT	917
QY	1159	AATCGATCGATTAATAATTTATGTTATATACAGATGAAGAGATTTTAAAGCCAGCGCA	1218
Db	918	KTTWTDDDDMWKARDKRWAAKADGAMWRPADDMWATATKDDGWDKMGGRKGRGK	977
QY	1219	GAGTCAAAATTTCTCGAGAAAAATGCTTAACACCATTCAGTGGATGGAGGAGAGAA	1278
Db	978	DKRRNDKRTGKTGKDDDDDKXTWTDNMWMTRTKWDNMWMDGRRGRMTRRKMGAMWRAD	1037
QY	1279	AAAAATAGTCGATTAAGACGCTTCTTGGAAGTAGAATAAAGAACCCGGGCGTTGGCT	1338
Db	1038	AAWADTDGKDTFTADKRRKRTDPTTKRGGDDMRKDRKDRKKGDDKTKKDATTWDDD	1097
QY	1339	AAT 1341	
Db	1098	ARD 1100	
RESULT	9		
CNSD03F8/c			
LOCUS	CNSD03F8	797 bp	DNA
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #		
	BACR08C03 of RPCI-98 library from Drosophila melanogaster (fruit		
	fly), genomic survey sequence.		
ACCESSION	AL064634		
VERSION	AI064634.1	GI:941986	
KEYWORDS	GSS.		
ORGANISM	fruit fly		
SOURCE	Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
AUTHORS	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
TITLE	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
JOURNAL	1 (bases 1 to 797)		
COMMENT	Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm . Location/Qualifiers 1..797 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_id="RPCI-98" /clone="BACR08C03" /note="end : TET3"		
FEATURES	source		
BASE COUNT	170 a	78 c	89 g 133 t 327 others
ORIGIN			
Query Match	2.7%	Score 41.6;	DB 122; Length 797;

Best Local Similarity 14.08; Pred. No. 1.4;
Matches 52; Conservative 160; Mismatches 159; Indels 0; Gaps 0;

[illegible]

RESULT	10
CNS0141M/c	
LOCUS	CNS0141M 1101 bp DNA
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC
	BACN11K19 of DrosBAC library from Drosophila melanogaster (fruit
	fly), genomic survey sequence.
ACCESSION	AL103598
VERSION	AL103598.1 GI:5615209

BASE COUNT	287 a	211 c	207 g	335 t	61 others
ORIGIN					
Query Match	2.7%; Score 41.6; DB 122; Length 1101;				

Best Local Similarity 45.7%; Pred. No. 1.5;
Matches .53; Conservative 23; Mismatches 40; Indels 0; Gaps 0;

QY	1234	GAGAAATGCGCTCAAAACACACATTCAGCTGATGGGACGACGAAAAAAATAATGCTCCGATA	1293
Db	1088	GCGCAAAAMCCCMSCSCCCCTTSCATGGGGGGCGDAAGSGSRAAAAAAAATTCGGAR	1029
QY	1294	AAGGACTTCTTGGAGATGAGATATAAAAGAACCGGGCGCTTGGGTATCTGAAAA	1349
Db	1028	HWRRGGGGRGGRAMCSRAAAAARWRARATTTTMCWTGGKDDTTTGGAAAA	973

RESULT	11
LOCUS	AQ258723
DEFINITION	npbb0021615r CUGI Rice BAC library Oryza sativa genomic clone
ACCESSION	npbb0021615r, genomic survey sequence.
	AQ258723

BASE COUNT	280 a	171 c	126 g	214 t
ORIGIN				

Wed Aug 23 11:46:10 2000

Search completed: August 22, 2000, 15:10:30
Job time: 828 sec

us-08-699-716a-1.rst

THIS PAGE BLANK (USPTO)

Page 10